PROBING THE SPATIAL AND FUNCTIONAL ORGANIZATION OF NETWORK CONNECTIVITY IN A SHORT-TERM MEMORY NEURAL CIRCUIT

A Dissertation

Presented to the Faculty

of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kayvon Patrick Daie

January 2014

PROBING THE SPATIAL AND FUNCTIONAL ORGANIZATION OF NETWORK CONNECTIVITY IN A SHORT-TERM MEMORY NEURAL CIRCUIT

Kayvon Patrick Daie, Ph.D.

Cornell University 2014

Experimental and theoretical work has suggested that persistent activity in short-term memory circuits is supported by excitatory recurrent connections. However, the organization of these connections remains difficult to accurately specify due to the typically low-dimensional nature of the recorded activity. To address this problem we use two-photon calcium imaging in the oculomotor integrator of larval zebrafish to study network activity under several different experimental conditions. By studying the network under multiple behavioral contexts, we have found that, unlike traditional "line attractor" models of connectivity, the network can support activity that persists for tens of seconds along two distinct dimensions which are accessed in a context-dependent manner. When put into a modeling framework, these dynamics put important constraints on the distribution of potential connectivity structures. To further refine our estimates of connectivity, we have used electrical microstimulations

to explore the dynamics along an increasing number of state-space dimensions. We find that the observed dynamics are most consistent with a picture in which connections are strongest between neighboring neurons.

Biographical Sketch

Kayvon grew up in Fountain Hills, Arizona and attended Fountain Hills High School. He then attended Arizona State University where he majored in Physics. Upon Graduating in 2006, he went on to pursue a Ph.D. in the Department of Physics at Cornell University in the Lab of Emre Aksay at Weill Cornell Medical College.

To Kate and KJ

Acknowledgements

During my time in graduate school, I have had the great privilege of collaborating with a number of excellent scientists, without whom the work described in this thesis would not have been possible.

First, I would like to thank my advisor Emre Aksay. As a member of the first generation of students in his lab, it has been a great pleasure to learn each of the steps involved in the design, execution and reporting of a good experiment directly from such a talented scientist. I am grateful that I was able to share this experience with fellow members of the Aksay lab Melanie Lee, Sherika Sylvester, Alex Ramirez and Charlotte Grove.

I have had the great fortune to participate in fruitful collaborations with the "integrator group" of Andrew Miri, Dimitry Fisher, Mark Goldman and David Tank from whom I not only learned the intricacies of this beautiful model system, but also a great deal about how to do neuroscience. In addition, it has been a pleasure collaborating with Ari Arrenberg, Ashwin Vishwanathan and Sebastian Seung.

During my first years of graduate school I was fortunate to learn a great deal of physics and the basics of doing good science from a number of great mentors and friends including my physics classmates: Ben Machta, Debo Oloasebikan, Srivatasan Chakram, Vikram Gadagkar, Phil Kidd and Sebastian Stolzenberg; as well as the Park

group: Mark Levendorf, Wei Tsen, Carlos Ruiz-Vargas, Lihong Herman, and Jiwoong Park; and my thesis committee: Itai Cohen and Tomás Arias. I would also like to thank the administrative staffs in both Ithaca and New York, including Kacey Bray Acquilano, for helping me navigate the complexities that go along with being a student *in Absentia*; and Carl Franck for exposing me to a constant enthusiasm for scientific discovery.

I would also like to acknowledge my family for their unwavering support of my journey to become a scientist; including my mom, dad and sister, Roya. Most importantly, I would like to thank my wife, Kate for her patience and willingness to move with me to Ithaca and New York City.

Table of Contents

List of Figures

Chapter 1 Introduction

1.1 Overview

The brain derives much of its computational capacity from connections between neurons. While much is known about the physiology of single neurons and about the role played by large scale (~100,000 neurons) brain regions in the control cognition and behavior, much less is known about the nature of the network-wide connectivity and how these connections relate to the activity observed at the single cell level. The advent of new technologies[1,2] for recording the electrical properties of entire populations of neurons simultaneously has led to immense growth in neuroscientists' ability to study the function of these neuronal microcircuits. Recent years have seen the establishment of a large-scale effort within the neuroscience community to develop new technologies capable of recording from even greater numbers of cells at increasing spatio-temporal resolution[3–5]. These methods should provide new insights into the question of *what* the brain is doing during the control of behavior and cognition. Combining these large-scale recordings of network dynamics with recently developed methods[6,7] employing electron microscopy to observe physical connections between networks of neurons should provide insight into *how* the brain generates the observed signals. However, a complimentary approach would

be required to understand the full pattern of connectivity as these techniques are limited in their ability to specify the actual strength of the interactions between the imaged cells. The aim of this thesis is to develop one such complimentary method, and to use it to gain insight into the connectivity underlying the storage of short-term memories.

A number of studies[8–11] have used measurement of electrical activity in networks of neurons to estimate the strength of all connections within the network. Each of these studies succeeded in identifying connectivity structures consistent with the measured activity, but each was limited in its ability to identify a unique structure. This limitation arises because the cells in the network tend to exhibit highly correlated patterns of activity. These correlations lead the regression-based techniques that are used for determining the connectivity to be rank-deficient. Such rank-deficiency, or sloppiness, has been identified as a feature present in modeling of many networks in systems biology[12].

An intuitive understanding of this sloppiness can be gained from a dynamical systems perspective. In this framework the activity of the network at a given instant in time is represented as a point in an N-dimensional space in which each axis represents the amplitude of the activity of a single neuron. The dynamics of the system can then be described by a trajectory though this space. From this perspective, interactions between neurons can be thought of as forming an energy landscape throughout state-

space which determines the full set of allowable trajectories[13–15], and so measurement of the trajectories at all points in state space should provide a complete understanding of the underlying connectivity. The sloppiness referred to in the previous paragraph is a result of the tendency of state space trajectories of neural circuits to be confined to an m-dimensional sub-volume of state-space where $m << N$ [9,16–19]. An m-dimensional trajectory yields no information regarding the connections responsible for shaping the energy landscape in the remaining N-m dimensions.

Low-dimensionality is an especially striking quality of networks responsible for storing short-term memories. The characteristic feature of these networks is an increase in electrical activity during stimulus presentation, followed by persistent activity that outlasts the termination of the stimulus[20,21]. Memories of stimulus amplitude are typically stored by the amplitude of firing along a small number of dimensions in state space[9,22]. It has been proposed in theoretical studies[15,23] that the connections in short-term memory networks are structured to support persistent activity along only a small number of dimensions, while activity along the remaining dimensions decays away quickly. The memory network for which the largest body of experimental support for this "attractor network" picture exists is the oculomotor integrator.

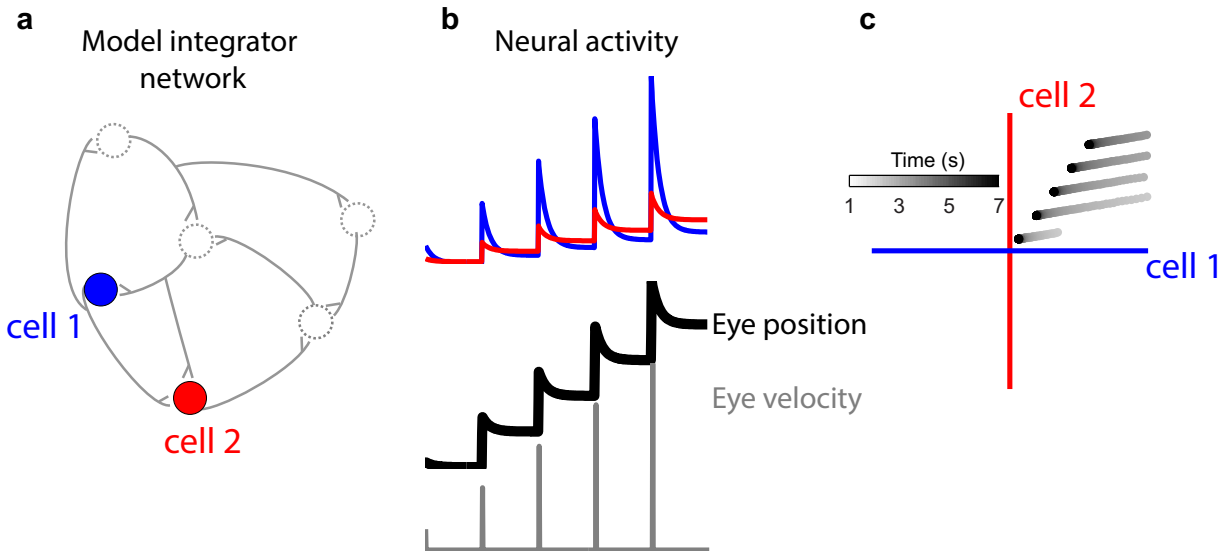


Figure 1.1: a.) Schematic illustrating connectivity of integrator neurons. b.) Electrical activity of two of the network's neurons (top) is correlated with eye position (bottom; black). c.) State space trajectory in the cell 1-cell 2 plane. Each trajectory represents the firing rate of these two cells during a fixation at a different eye position. The firing rates decay onto a line with the location along the line coding the position of the eyes.

The oculomotor integrator is the brainstem network responsible for stabilizing the position of the eyes. Movement of the eyes is accomplished by either internally-generated or sensory-evoked eye velocity signals. After the termination of movement commands, the eye muscles require a constant stream of input in order to maintain the new position. In the absence of persistent inputs the eyes would relax back to their resting position. This persistent activity is produced by the integrator whose role is to generate persistent eye position signals in response to transient eye velocity inputs. When the eyes are held still at positions away from the center of gaze, the amplitude of persistent activity in integrator neurons, *r*, is approximately related to the angular eye position by the relationship $r = kE + r_0$ [24] (Figure 1.1b). Because the activity

can be approximately parameterized by a single variable, the activity of the network exists mainly along a single dimension in state space Figure 1.1c. Furthermore, a number of complimentary experiments have suggested that connections between integrator neurons represent the dominant mechanism underlying the maintenance of persistent activity in the integrator[25–27]. For example, Aksay *et. al.*[27] found that silencing a fraction of the cells in the network with the drug lidocaine resulted in a loss of persistent activity the rest of the network, suggesting that the feedback from the inactivated neurons was important in the production of persistent activity. However, because integrator activity is confined to a small number of dimensions in state space the full connectivity profile remains unknown.

Early work proposed that the integrator functioned essentially as a "line attractor", in which activity decayed quickly from all points in state space onto the line described above. These models have been difficult to disprove because of the lack of activity recorded away from the dominant line. In light of this, the line attractor represents the most reasonable description of the unexplored dimensions since the activity of a neuron, by default, is leaky. If a network's sole purpose is to generate persistent activity along a single dimension, there would appear to be no benefit for it to tune its connections to exhibit varied dynamics along any other dimensions. However, the integrator, like other brain areas, receives inputs from, and provides outputs to, a number of distinct brain regions, each with potentially different

processing requirements. In Chapter 3 we explore the possibility that the network supports non-trivial dynamics along modes away from the dominant line by characterizing network activity under multiple behavioral contexts. We find that the network is capable of supporting activity that persists on the ten seconds time scale along two dimensions, one accessed during internally-generated quick eye movements, and another accessed by visually-evoked slow eye movements. These context-dependent dynamics may prove useful in optimally controlling behavior.

The results of Chapter 3 also reveal potential insights regarding the connectivity of the network. At the single-cell level, we find that cells tend to have dynamics that are similar to their neighbors. Intuitively this finding suggests a connectivity profile in which neurons connect more strongly with their neighbors. In Chapter 4 we use mathematical modeling of network activity to show that this spatial organization is indeed consistent with the observed activity. However, because of the low-dimensional nature of the observed dynamics, this organization represents only one of many structures consistent with the data. The rest of Chapter 4 is devoted to quantifying the complete space of networks consistent with the activity in an attempt to identify general features consistent among all networks. Ultimately, this analysis yields ambiguous results regarding the spatial organization of the network's connections.

The context-dependent dynamics described in Chapter 3 underline the importance of studying activity along all dimensions state space in order to fully understand connectivity. To record activity along a greater number of dimensions, in Chapter 5 we describe the use of electrical microstimulations to perturb activity in the network along dimensions not active during behavior. When considered together with behaviorally-related dynamics, these stimulated responses provide strong support for the hypothesis provoked by the spatial organization of activity in Chapter 3 that connections are strongest between neighboring neurons.

Chapter 2 Background

2.1 Abstract

In this chapter I outline the method used in Chapters 3-5 help identify the connectivity structure most likely underlying the generation of persistent activity in the integrator. Before outlining this method, I provide a brief outline of the physics relevant to understanding the function of neural circuits.

2.2 Biophysics of neural circuits

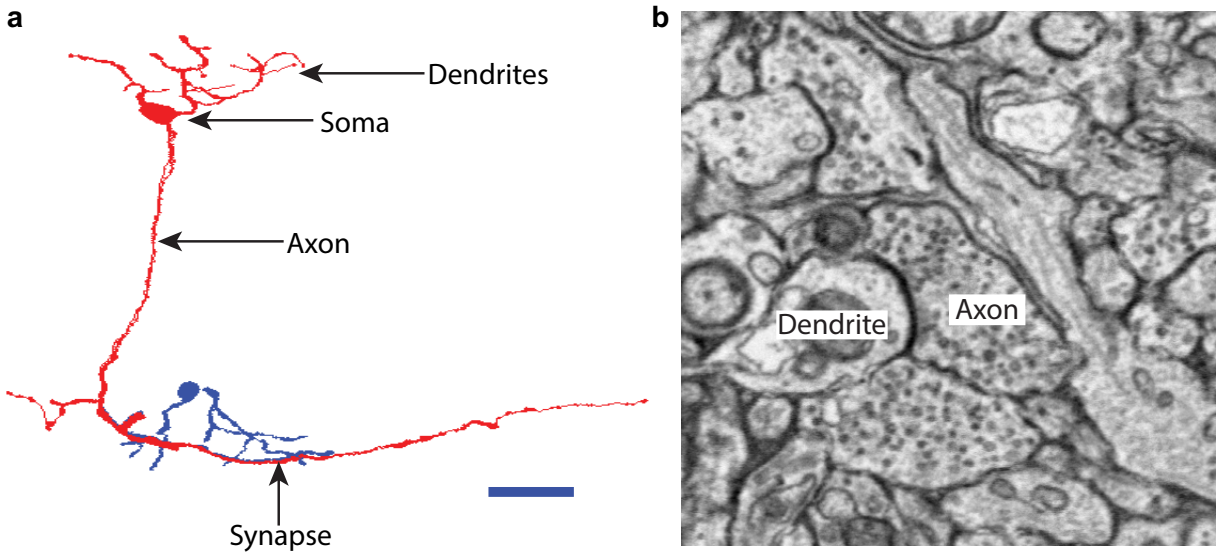


Figure 2.1: Illustration of the main structural features of a neuron. a.) Neuronal morphology visualized using two-photon fluorescence microscopy. The overlap between the axon of the red cell and the dendrites of the blue cell is an example of a putative synaptic contact. Scale bar represents 20 μm. b.) Electron micrograph of region in zebrafish hindbrain dense in axons and dendrites. At this resolution synapses can be visualized as shown. The dark spots visible in the axon are neurotransmitter-containing vesicles. Scale bar represents 0.4 μm.

The brain uses electrical signals to control behavior and cognition. The substrate used to transmit these signals is the voltage difference between the inside and outside of the neuronal membrane, or the membrane potential. This voltage difference is established via concentration gradients in ionic species established by proteins present in the neuron's lipid membrane. Signaling of the present state of a neuron's membrane potential is facilitated by the generation of rapid transient fluctuations in membrane potential—termed action potentials—mediated by several classes of voltage gated ion channels. The channels underlying the generation of

action potentials are distributed along the surface of the neuron's output process, the axon. Once an action potential is generated within the cell body it is able to travel down the axon. The arrival of an action potential at an axon terminal triggers the release of neurotransmitters. These neurotransmitter chemicals can interact with receptors on the input processes, or dendrites, of neighboring neurons, an interaction which can have the effect of opening any one of a large variety of ion channels. Such an interaction between two neurons is referred to as a synapse (Figure 2.1) and is the most common and well-studied way in which neurons can transmit signals to one another.

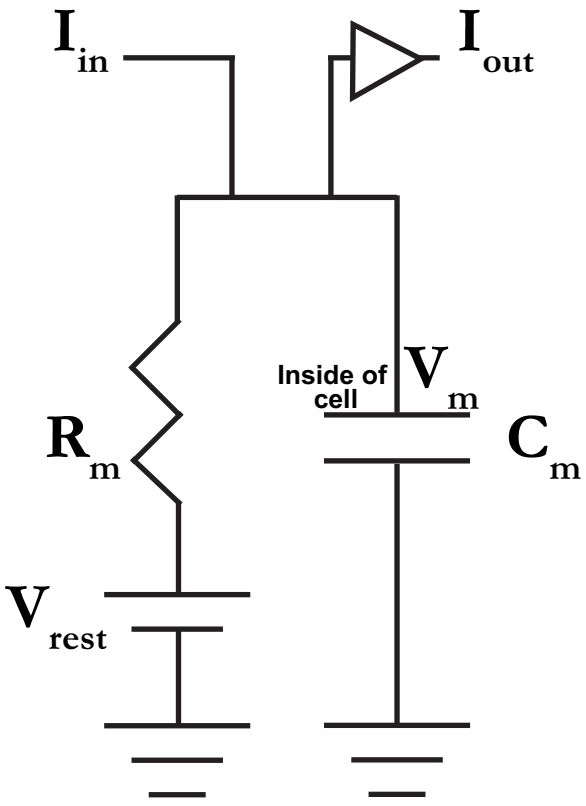


Figure 2.2: Circuit model of the electrical properties of a neuron.

The membrane potential can be modeled in the simple circuit shown in Figure 2.2. The capacitor represents the capacitive properties of the membrane itself while the resistor represents the channels in the membrane through which ions flow to restore the neuron back towards its resting state. Given these properties, the membrane potential can be modeled by the following equation

$$\tau_m \frac{dV}{dt} = -V + IR_m \tag{2.1}$$

where the time constant τ_m is simply the product of the membrane resistance R_m and capacitance C_m and is typically of order 10 milliseconds. From this equation we see that the membrane voltage is simply a low-pass filtered version of the current into the neuron. Furthermore, because there exists an approximately linear relationship between membrane potential and the rate at which the neuron generates action potentials[25], we can model the neuron's firing rate, r, using the equation

$$\tau_m \frac{dr}{dt} = -r + IR_m \tag{2.2}$$

The synaptic current into the post-synaptic neuron is itself a function of the firing rate of the pre-synaptic neuron and therefore the equation describing the interaction of an entire network of neurons takes the form

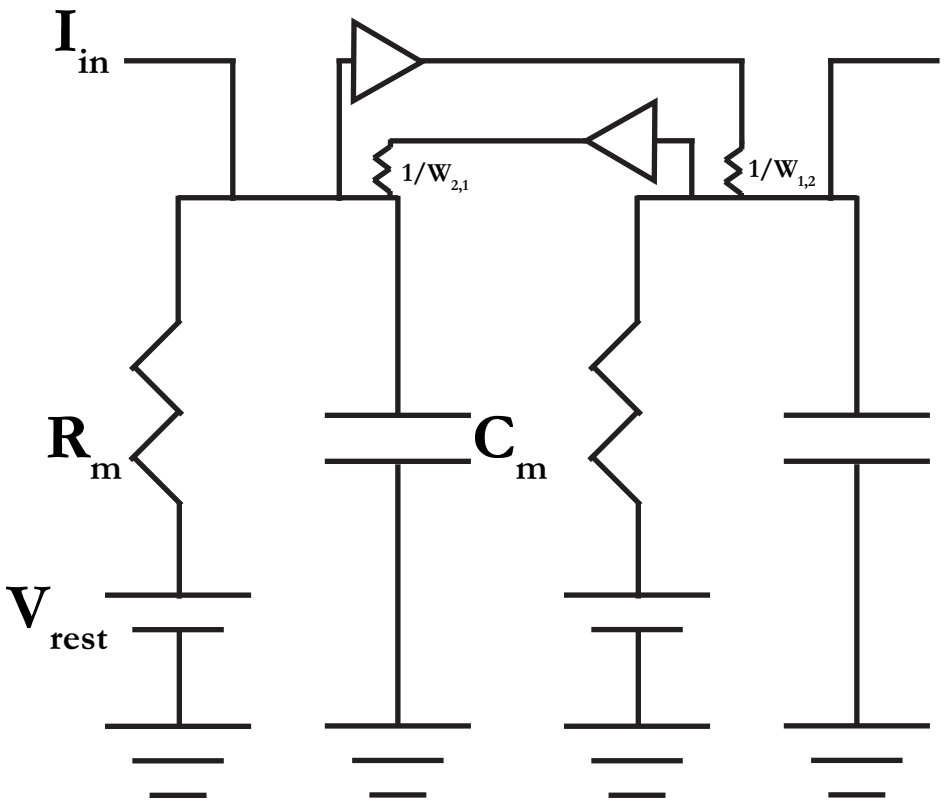


Figure 2.3: Equivalent electrical circuit of a two-neuron network.

$$\tau_m \frac{dr_i}{dt} = -r_i + \sum_j f(r_j) W_{j,i} + h_i \qquad 2.3$$

where $W_{j,i}$ is the strength of the connection from cell j onto cell i and h_i is the input to cell i from outside of the network. $f(r)$ is the functional relationship between the synaptic current and the presynaptic firing rate. $f(r)$ is most accurately modeled as a saturating function such as a sigmoid[8,13,28], however, over a large portion of the firing rate range, this relationship can be well-approximated as linear.

Given that the activity of a neural circuit can be modeled in this simple framework, these circuits can be understood using the techniques developed in the field of dynamical systems. In this framework, the dynamics that can be generated by the system are determined by the nature of the interactions between the constituent elements, in this case its neurons. Here we seek to characterize the dynamics of the oculomotor integrator during the control of behavior, and develop complimentary experimental and theoretical techniques for indentifying the patterns of interactions underlying the generation of persistent activity.

2.3 Modeling integrator network dynamics

To model the response of the integrator during a quick eye movement, or a saccade, we can set the external input in equation 2.3 equal to $\mathbf{h}(t) = \mathbf{h}\delta(t)$. The delta function is meant to represent the eye velocity input to the network during a saccade. Given this input, the firing rate of the i^{th} cell is

$$r_i(t) = \sum_j c_{0,j} e^{-t/\tau_j} V_{i,j} \qquad 2.4$$

where $\mathbf{V}$ is a matrix whose columns are the eigenvectors of $\mathbf{W}$ and the i^{th} time constant is related to the i^{th} eigenvalue by

$$\tau_j = \frac{\tau_m}{1-\lambda_j} \qquad 2.5$$

The initial value along a given eigenmode, $c_{0,j}$ is determined by the degree of overlap between the input vector **h** and the j^{th} eigenvector. Equation 2.5 illustrates the way in which the network's connectivity can be used to lengthen the time constant of persistent firing. For example, if $\lambda_1 = 1$, the time constant along the first eigenvector would be infinite and a pulse of input would be integrated perfectly by the network into a step of output. Using this framework, the simplest model of an integrator is one in which

$$\lambda_i = \begin{cases} 1, & i = 1 \\ 0, & i \neq 1 \end{cases} \qquad 2.6$$

As mentioned in the introduction, such a network is referred to as a line attractor since all activity decays quickly onto a single line, defined by the i^{th} eigenvector, in state space.

2.4 Illustrating sloppiness in an integrator network

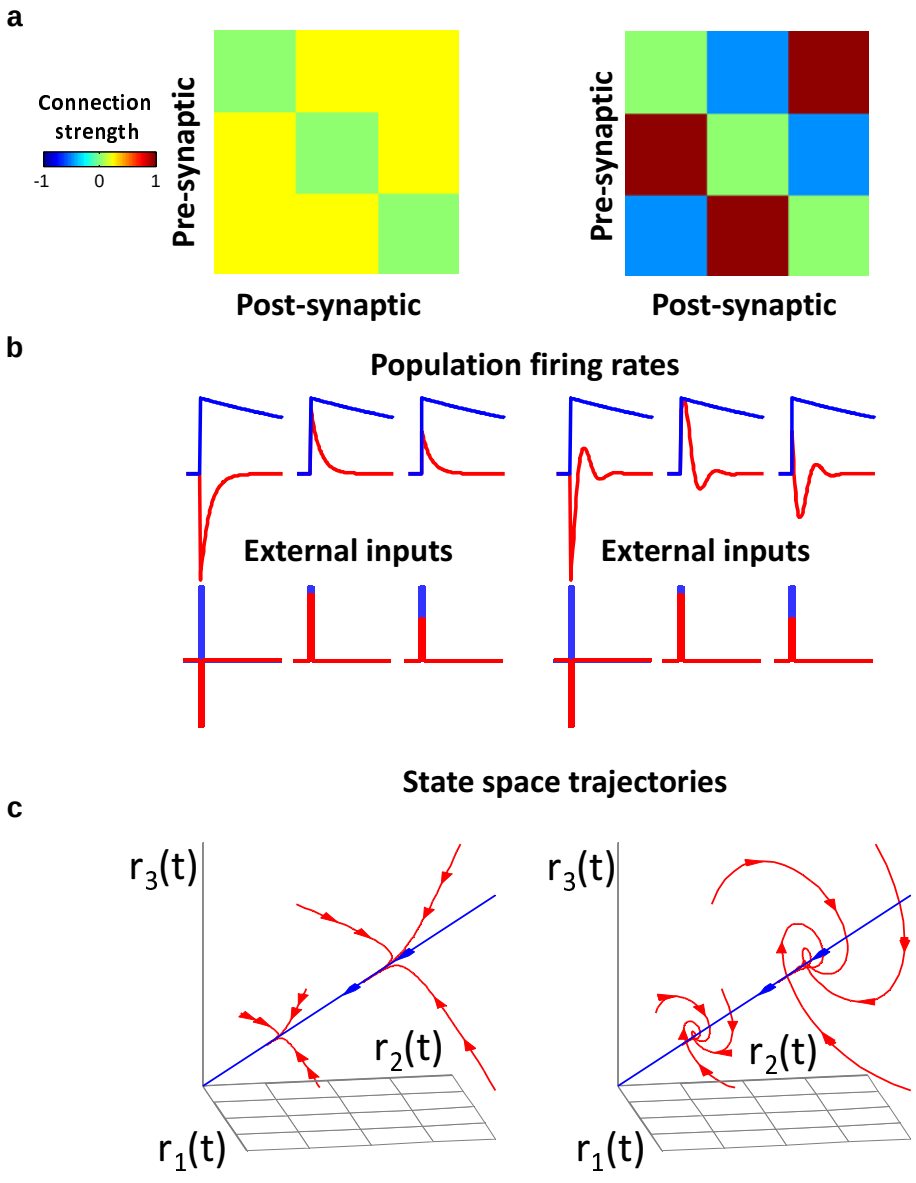


Figure 2.4: a.) Connectivity matrices. b.) Firing rates of integrator neurons (top) in response to a pulse of input (bottom). c.) State space trajectories.

The following sub-section provides a simplified example of the difficulty encountered when attempting to identify the connectivity of an integrator network from recordings of the firing rates of all cells in the network. In this example, we simulate activity in the two networks whose connectivity structures are represented in the matrices in Figure 2.4-a. These two networks have different connectivity structures, but both support persistent firing with a 10 second time constant along the eigenvector ({1,1,1}). If velocity inputs to these two networks were aligned perfectly with this eigenvector (Figure 2.4b; blue traces, bottom), the two networks would show identical patterns of activity (Figure 2.4b; blue traces, top). An experimenter encountering these dynamics would be unable to determine which of these two networks, or any network with an integrating mode along the {1,1,1} direction, represented the true connection structure.

The reason for this ambiguity is the lack of information along the modes orthogonal to the {1,1,1} vector. While the two networks shown here exhibit identical activity along this direction, they each support distinct dynamics along their other two eigenvectors. Therefore, in order to accurately specify the underlying connectivity, it would be necessary to observe the activity along all three of the networks eigenvectors. In Figure 2.4b the red traces show the result of exciting the two networks with an input vector orthogonal to {1,1,1}. Unlike the {1,1,1} stimulation, the orthogonal stimulation produces different patterns of activity in the two networks.

State space trajectories in the space away from the vector {1,1,1} are shown in Figure 2.4c further illustrating that the dynamics along the modes orthogonal to {1,1,1} are different for these two networks. As we will illustrate in the next section, specification of the exact connectivity structure can be achieved by characterization of the dynamics along all dimensions in firing rate space.

2.5 Using random stimulations of network activity to determine connectivity

The discussion in the previous section illustrated that specification of the connectivity from recordings of population dynamics requires high dimensional dynamics. As we will show in Chapter 3, the dynamics in the integrator during the control of eye movements, like many other systems, tend to be low-dimensional. However, increased dimensionality can be accomplished by manually perturbing the activity of the network into different modes. Temporally precise perturbation of neural activity can be accomplished either electrically[2] or optically by inserting light-gated ion channels into the circuit of interest[29,30].

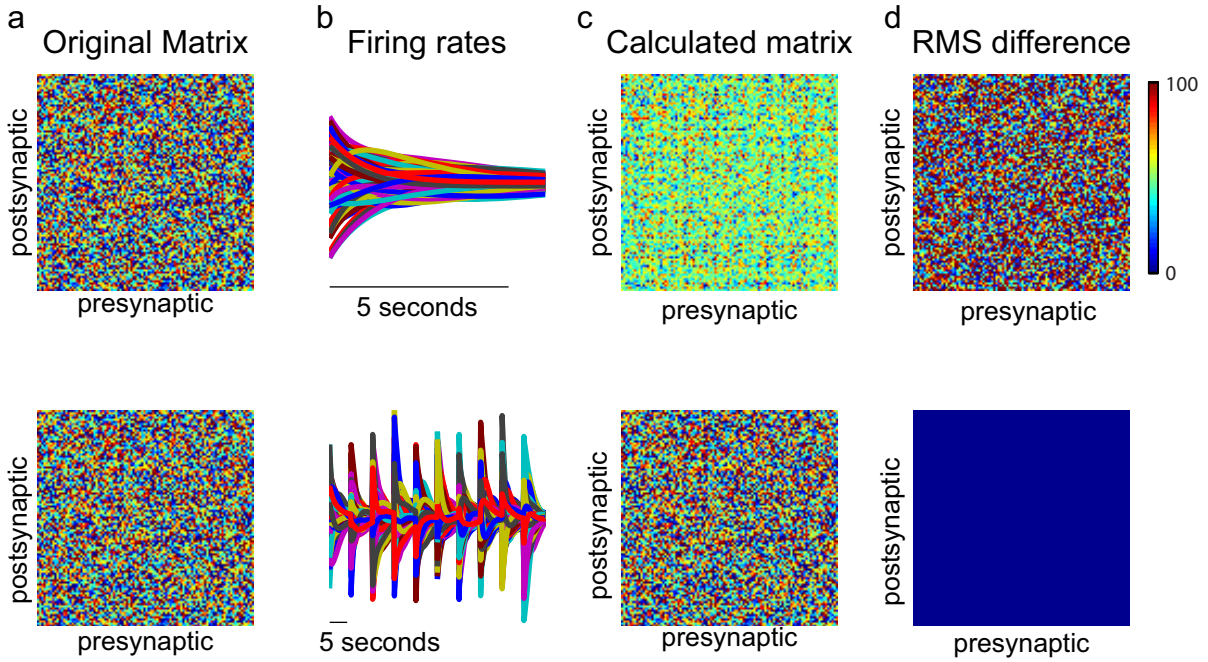


Figure 2.5: Using random perturbations of electrical activity to determine network connectivity. a.) initial random connectivity matrix. b.) Activity produced with a pulse of input with randomly specified amplitude for each cell (top) and ten such randomly stimulated patterns (bottom). c.) Connectivity matrix that fits responses in b. d.) RMS difference between initial and best-fit matrix.

The increased accuracy in estimates of connectivity gained from random perturbations is illustrated here using a simple simulation (Figure 2.5). In this simulation we start with a randomly connected matrix (Figure 2.5-a). The network is then provided with a brief pulse of input in which the input vector (strength of the pulse onto each cell) is randomly specified resulting in the activity pattern shown in Figure 2.5b (top). We can then use the activity following the termination of the input to solve for the connectivity matrix using linear regression, a procedure which we

describe in greater detail in Chapter 4; the solution is shown in the top panel of Figure 2.5c. Because of the low dimensional nature of the responses in Figure 2.5b the solution does not match the initial connectivity matrix. If we stimulate the network using multiple random patterns as shown in the bottom panel of Figure 2.5b, and use all of these independent responses together to solve for the connectivity, the accuracy of the solutions increases until eventually we reach the correct answer.

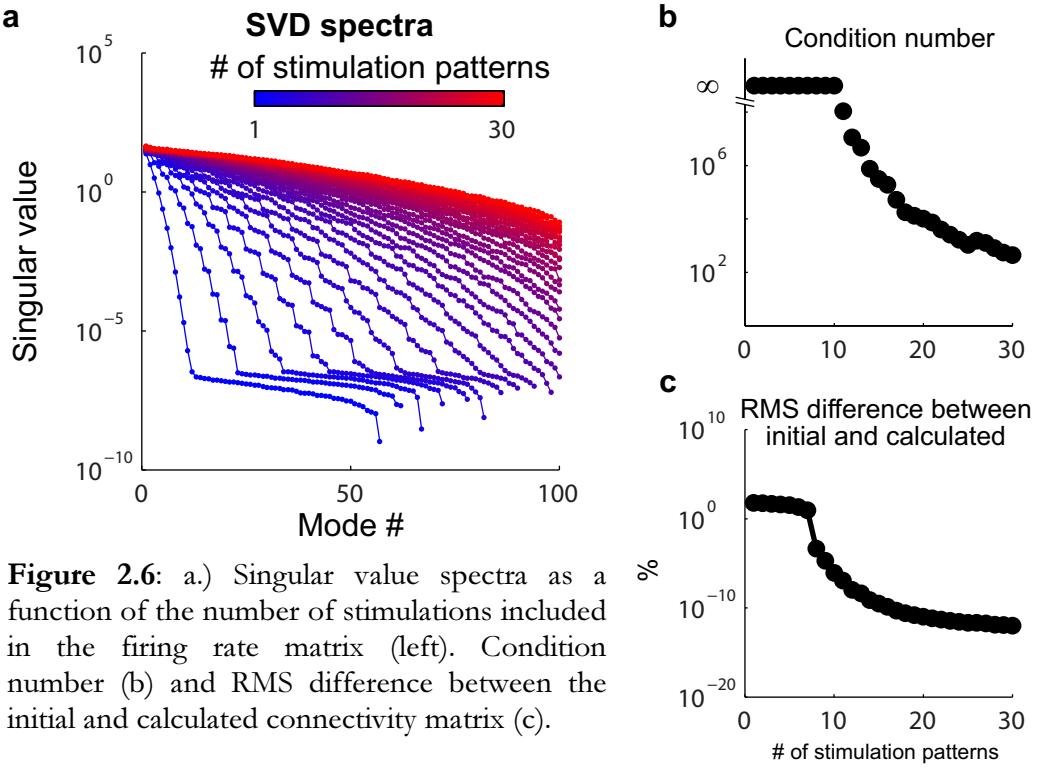


Figure 2.6: a.) Singular value spectra as a function of the number of stimulations included in the firing rate matrix (left). Condition number (b) and RMS difference between the initial and calculated connectivity matrix (c).

The convergence onto the correct connectivity matrix with increased independent stimulations is further illustrated in Figure 2.6. The increased accuracy that results from adding new patterns of stimulation arises from the increased dimensionality of the firing rate matrix. This increase in dimensionality is best

quantified via the singular value decomposition (SVD, Chapter 3). The SVD is a technique which can be used to identify the dimensions (or modes) in firing rate space with the largest root-sum-squared (rss) amplitude (i.e. variance relative to zero). The i^{th} singular value is a measure of the rss amplitude of the signal along the i^{th} mode. SVD spectra for firing rate matrices as a function of the number of independent stimulations are shown in Figure 2.6a. We see that as more random patterns of stimulation are added (more red) the ratio between the largest and smallest singular values increases indicating increased dimensionality of the data. The ratio between the largest and smallest singular value is referred to as the condition number and is shown as a function of the number of stimulation patterns in Figure 2.6b. We see that when fewer than 11 stimulations are included, the condition number is infinite. Once at least 11 stimulations are included all singular values are greater than zero—and hence the condition number is less than infinity—which means that the activity of the network is present along all dimensions in state space. The plot in Figure 2.6c shows the RMS difference between the solution and the initial matrix. Here we see that once the network has been excited along all dimensions (# of stimulations is greater than 11) the difference between the initial and calculated matrices drops off rapidly.

2.6 Overview

In Chapter 4 we find that the dynamics during eye movements characterized in Chapter 3 are consistent with connectivity structures with a wide variability in their spatial organization. For example, we find networks in which the strongest connections are between nearby neurons, but also networks in which the strongest connections are between far away neurons, are capable of fitting the dynamics during eye movements equally well. To distinguish between these vastly different structures, in Chapter 5 we apply the approach described in the previous section. As expected we find that each time a new pattern of activity is generated via electrical stimulations, we obtain successively more refined estimates of connectivity until finally we find that the most likely connectivity structure is one in which cells connect preferentially to their neighbors.

Chapter 3 Spatial Patterns of Persistent Neural Activity Vary with the Behavioral Context of Short-Term Memory[1]

3.1 Abstract

A short-term memory can be evoked by different inputs and control varied targets in different behavioral contexts. To address the circuit mechanisms underlying context-dependent memory function, we determined through optical imaging how a memory is encoded at the whole-network level in two behavioral settings. Persistent neural activity maintaining a memory of desired eye position was imaged throughout the oculomotor integrator after fast (saccadic) or slow (optokinetic) eye movements. While eye position was encoded by the amplitude of network activity, the spatial patterns of firing were context-dependent: cells located caudally (closest to the tail) were most persistent following saccadic input, whereas cells located rostrally (closest to the head) were most persistent following optokinetic input. To explain these data, we computationally identified four independent modes of network activity and showed that saccadic and optokinetic signals differentially access these modes. These results show how a circuit can simultaneously encode memory value and behavioral context, respectively, in its amplitude and spatial pattern of persistent firing.

[1] This chapter is adapted from a manuscript that is in review, K. Daie, M. Goldman and E. Aksay, 2013.

3.2 Introduction

Short-term memory circuits can accumulate information arising from different senses or brain regions and can provide outputs to different targets to meet the varied demands that arise within different behavioral contexts[31]. The accumulation and storage of information in short-term memory is a critical component of a range of behaviors including executive control, decision-making, spatial navigation, and motor control[32,33]. For example, during delayed-comparison working memory tasks, cells in the prefrontal cortex (PFC) accumulate briefly presented signals and subsequently store relevant information through persistently maintained firing[34,35]. During cognitive decision making, the firing of cells in the lateral intraparietal region (LIP) encodes a memory of accumulated evidence[36]. And, during oculomotor control, cells of the velocity-to-position neural integrator (VPNI) temporally accumulate brief velocity-encoding inputs and store the resulting output as persistent firing rates encoding desired eye position[37,38,39]. In all of these cases it remains unclear how short-term memory circuits exhibit the flexibility needed to generate the different input-output relationships required for context-dependent function.

A major challenge in addressing the neural basis of context-dependent processing during memory behavior is to visualize the activity of a memory circuit under different behavioral contexts. Here we overcome this challenge in the VPNI for

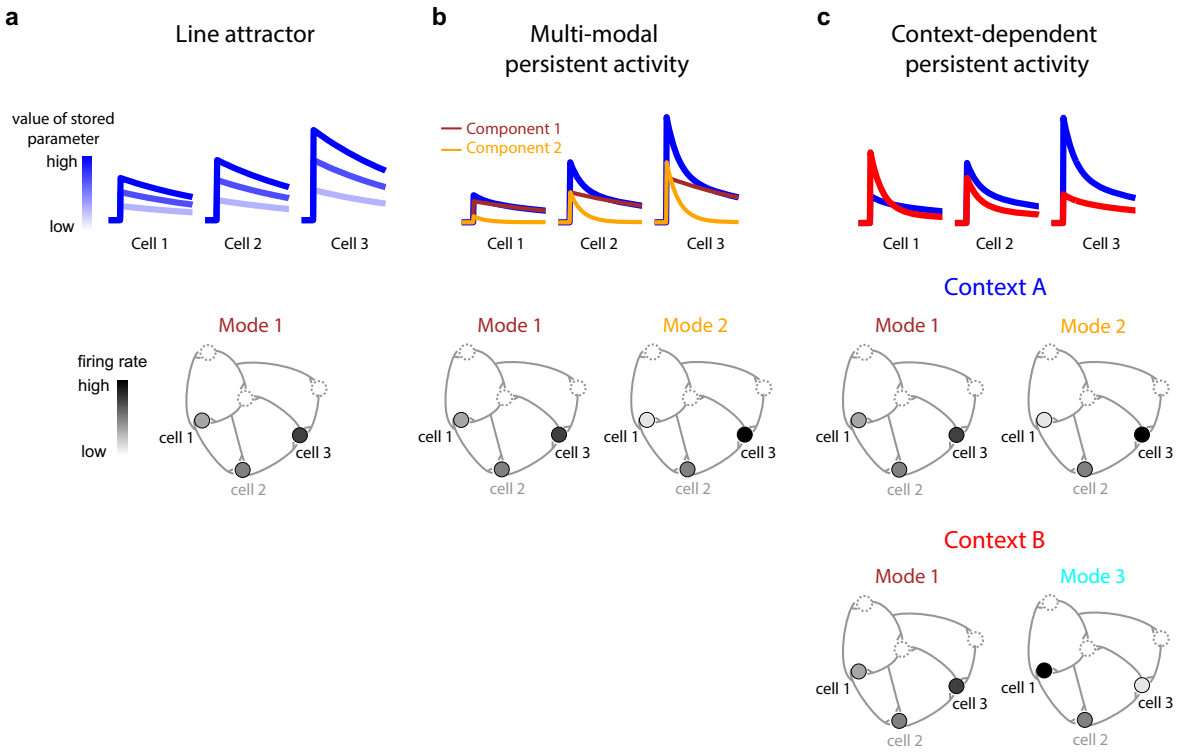


Figure 3.1: A potential solution for the challenge of short-term memory storage and temporal integration under different contexts. (a) Persistent activity for three cells (top) in a line-attractor network defined by a single pattern, or mode, of activity (bottom). The value of the memory variable is stored in the amplitude of activity (top) along this mode. The relative firing rates of the constitutive neurons in the memory network, which is a constant across different memory values, defines the mode (bottom). (b) Heterogeneity in the persistent activity of the same three cells (top) when two modes are active (bottom). The activity of each cell is given as a linear combination of the firing rate components associated with each mode. These two components are displayed in the red and yellow traces. (c) If the integrating network is capable of supporting multiple modes of activity, then different behavioral contexts can use different sets of modes. In this schematic, the two modes of panel b are accessed during one context (context A, top), and a different combination of modes are accessed in a second context (context B, bottom). At the single cell level (top), a signature of such context-dependent processing in a graded short-term memory circuit or integrator would be firing rates that differ across contexts in both amplitude and rate of drift back to baseline.

horizontal eye movements by using two-photon microscopy in the larval zebrafish to image the context-dependence of persistent activity at the whole-circuit scale. The VPNI mathematically integrates transient inputs encoding desired eye velocity from a

variety of upstream sources[24,40,41], and sends its output, the desired eye position, to multiple targets including the oculomotor nuclei, cerebellum, superior colliculus, and thalamus[42,43]. Although the essential function of this system, the conversion of velocity-encoding inputs to position-encoding outputs, is preserved across different oculomotor behaviors, the specific transformations that occur within the VPNI may be context-dependent in order to handle the particular requirements of different eye movement behaviors such as rapid gaze-shifting (saccades), or slow tracking of visual stimuli (optokinetic tracking), or compensating for head movements through the vestibuloocular reflex.

Recent findings on VPNI physiology offer important clues as to how context-dependent processing might occur in this system. Mechanistically, maintenance of information in the VPNI depends in large part on positive feedback generated through recurrent excitatory synaptic interactions[25,27]. Traditionally, it has been assumed[15,44,45] that this positive feedback tightly couples integrator neuron firing rates so that there exists, in all behavioral contexts, a single pattern of activity across the network specified by the relative firing rates of the constitutive neurons. That is, when the system responds to stimuli of different amplitudes, the overall level of persistent firing differs, but the pattern of activity is maintained. In the language of dynamical systems, this pattern is referred to as a *mode* of network activity, and systems with a single, stable persistent mode towards which the network's activity quickly relaxes are

referred to as line attractors (Figure 3.1a). However, while the idea that temporal integration is dominated by a single, line attractor mode of activity has been applicable to a range of experimental results[33,46], multi-unit studies have revealed considerable heterogeneity across the VPNI population[11,47]. This heterogeneity suggests the presence of two or more modes governing VPNI population dynamics, where each mode is specified by a different pattern of relative firing rates and is associated with a different timescale of persistent activity[11,47,48] (Figure 3.1b).

The presence of multiple modes in the VPNI could potentially offer a solution to the challenge of how short-term memory storage can be dynamically altered within different behavioral contexts (Figure 3.1c). First, because each mode can be accessed independently and effect a specific response in downstream targets, different modes (or combinations of modes) could be associated with different contexts. Second, because the modes represent dynamically organized groups of neurons established by a single set of synaptic connections within the circuit, they could be used to rapidly switch function as the contextual setting changes. Third, because modes can potentially be activated simultaneously, they could be used to coordinate multiple functions at the same time.

Here we determined whether different modes of activity in the VPNI are used for contextual processing. We used calcium imaging to record the activity of integrator neurons throughout the VPNI circuit during the storage of a short-term

memory of eye position in two different behavioral contexts, saccadic gaze-shifting and optokinetic tracking. We first determined what modes of activity are present in the VPNI and then examined if these modes are used in a context-specific manner. We expected that if different modes were used in a manner that varied with the behavioral context, then we would find that the time courses of individual neurons' firing rates should differ systematically in the two settings (Figure 3.1c).

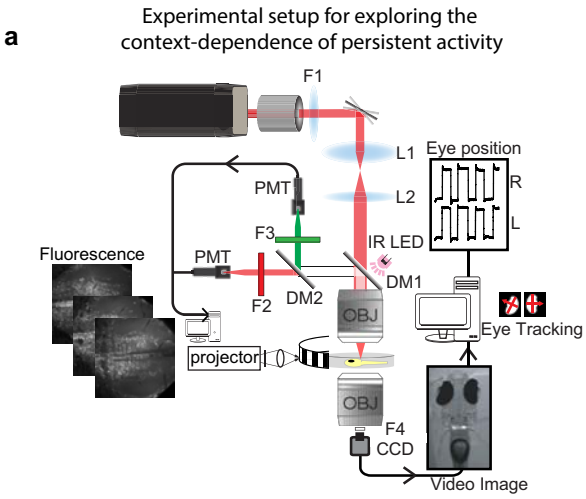
We show that while the overall magnitude of persistent firing represented *what* position value the eye should be held at, the pattern of persistent firing across the circuit represented *how* the eye position signals were initiated, i.e. the behavioral context. These patterns of persistent firing reflected four underlying modes of activity that were utilized in a context-dependent manner. Strikingly, a switch in behavioral context was coupled to a rapid reversal across the circuit in a spatial gradient of persistent activity. Using computational modeling, we also show how the excitatory synaptic interactions previously implicated in generating persistent activity in this system[25,27,8] can lead to the observed context-dependent firing patterns through the selective activation of different modes by different classes of inputs. Together, these results suggest a general framework by which to understand how short-term memory circuits function under different behavioral contexts.

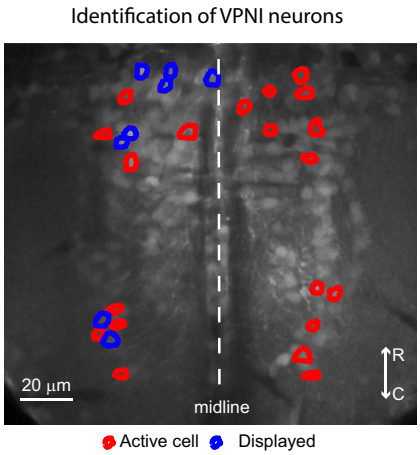
3.3 Results

To investigate how contextual information is represented during short-term memory, activity in the velocity-to-position neural integrator (VPNI) was assessed through recording and analysis of calcium-sensitive fluorescence responses during both spontaneous saccadic and optokinetic behavior. In the following, we identify for each of these contexts the patterns of firing underlying the maintenance of eye fixations following the cessation of external input, when the dynamics of eye position are dependent on persistent activity generated in the VPNI[49]. We then compare across behaviors the dynamics and spatial organization of firing patterns in the network to quantify the impact that context has on persistent activity. Finally, we construct a model of the VPNI[25] that incorporates a range of experimental results in this system[25–27] and explains how context-dependent processing of different short-term memory inputs can occur in a single circuit.

3.3.1 Persistent activity after saccadic and optokinetic stimulation

We measured the activity of neurons in the horizontal VPNI of larval zebrafish using two-photon calcium imaging (Figure 3.2a) and directly compared the dynamics of firing after saccadic eye movements to the dynamics after eye movements driven by the optokinetic reflex (OKR). Integrator neuron activity was imaged in horizontal





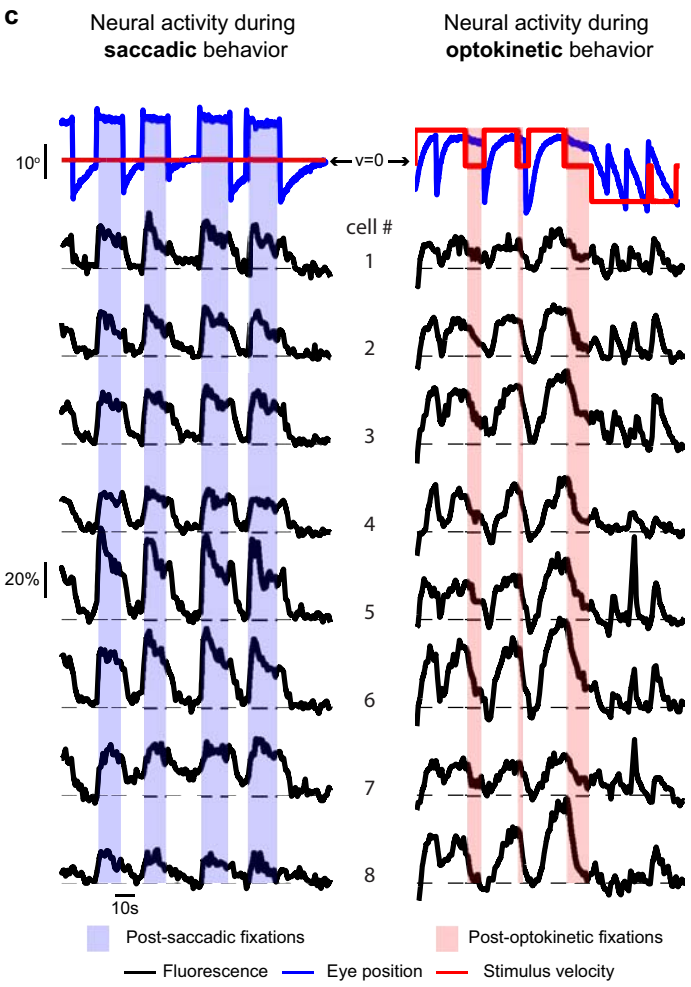


Figure 3.2: Experimental setup for exploring the context dependence of persistent activity. (a) Schematic of experimental set-up used for synchronous two-photon calcium imaging, behavioral imaging, and eye-tracking (DM: dichroic mirror, F: filter, L: lens). (b) Average image of one plane in the caudal hindbrain where VPNI neurons were located. The Mauthner axons are visible on either side of midline. Outlines indicate cells with eye position correlated activity; blue outlines indicate those that are plotted in (c). (c) Eye position and stimulus velocity (top), and fluorescence time series of individual cells (bottom) during saccadic and optokinetic eye movements. Dashed lines indicate the baseline level of fluorescence for each cell. Colored bars indicate post-stimulus fixation regions where gaze stability is dependent on persistent firing generated within the VPNI.

planes after bolus-delivery to the caudal hindbrain of Oregon-Green Bapta-1 AM (Figure 3.2b), a calcium-sensitive indicator which reliably reports firing rate changes in these cells[50]. VPNI somata were identified using a modified version of a previously described method to find pixel clusters in which fluorescence was significantly correlated to eye position[50]. In total, 262 integrator cells from 10 zebrafish were included in the analysis. All cells were recorded during both saccadic and optokinetic behavior.

Our experimental paradigm was designed to test if network-level activity in the VPNI differed with stimulus context. Activity was measured during fixations at ipsilateral, eccentric values of eye position whose maintenance are known to depend upon the firing of VPNI neurons both in the dark[50] and in the light[11]. Fixations were initiated in one of two ways. First, zebrafish spontaneously generated saccades to eccentric positions while viewing a stationary optokinetic stimulus (Figure 3.2c, left); these fixations depended upon temporal integration of velocity signals from the saccadic burst generator[51]. Second, zebrafish tracked transient movement of the optokinetic stimulus to eccentric positions, at which point stimulus movement was stopped (Figure 3.2c, right); these fixations depended upon temporal integration of velocity signals from pretectal pathways[52,53]. Transitions between experimental paradigms typically occurred in less than 10 seconds. Mean eye position during "post-saccadic fixation" was 7.3 ± 0.4^{o} (s.e.m), with nasally-directed drift of 0.26 deg/sec.

Mean eye position during "post-optokinetic fixation" was 7.1 ± 0.3^{o}, with nasally-directed drift of 0.34 deg/sec. These average eye positions did not differ significantly ($p=0.49$); furthermore, the average difference between left and right eye position during post-saccadic (0.42 ± 0.40^{o}) and post-optokinetic fixation (0.31 ± 0.55^{o}) also did not differ significantly ($p=0.88$). Thus, during both fixation types, the VPNI had to represent a similar value, namely at what position the eye was held, and it also may have represented contextual information, namely whether the fixation was initiated through saccadic or optokinetic stimulation. We expected that if the VPNI was able to retain information about context, then during fixation the patterns of persistent firing for the two behaviors should differ significantly (Figure 3.1c).

We next briefly describe the qualitative features of the neural activity during these two behaviors (Figure 3.2c) before performing a quantitative comparison in the following

sections. During ipsiversive saccades (those directed ipsilaterally), integrator cells

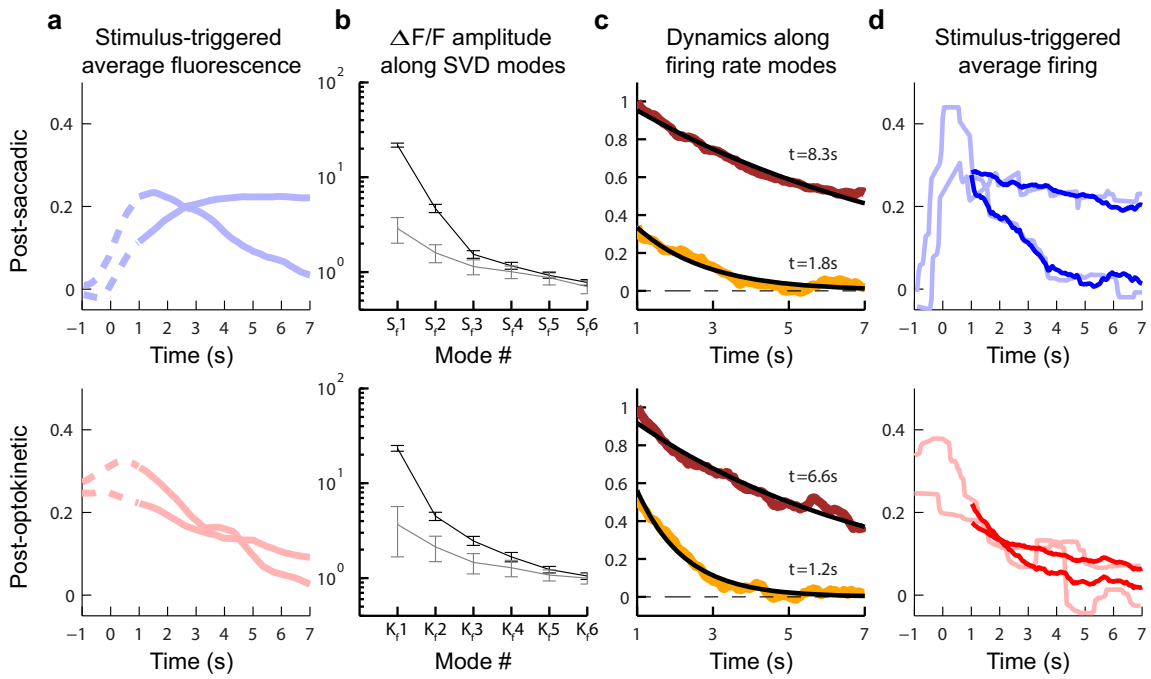


Figure 3.3: Mode analysis of responses during post-saccadic (top), or post-optokinetic (bottom) fixations. (a) Stimulus-triggered fluorescence responses (ΔF/F) for two cells during saccadic, and two other cells during optokinetic, eye movements. Time zero corresponds to the end of the stimulus. Solid lines represent the analyzed portion of the fixation period. (b) Singular values of the stimulus-averaged responses (black), which give the root-sum-square (rss) amplitudes of each fluorescence component. These are compared to the average rss amplitudes from noise (gray). Error bars give the 90% confidence intervals. (c) Projection of the population firing rates along the primary (dark red) and secondary (yellow) firing rate modes (S1, S2, K1 and K2), and their single exponential fits (black). (d) Deconvolution of average fluorescence data from a gives firing rates corrupted by noise (light lines, smoothed with a 1 second moving median filter, MATLAB medfilt1), while deconvolution after using the singular value decomposition to separate signal from noise gives firing rates accurately capturing the underlying dynamics (thick lines).

uniformly showed a rapid rise in activity. During the subsequent fixation (shaded blue), fluorescence levels either remained elevated relative to baseline (e.g. cell 4) or relaxed back towards initial values (e.g. cell 6). During contraversive saccades (those directed contralaterally), activity levels rapidly fell back to baseline if they were still

elevated at the end of the fixation. During constant ipsiversive movement of the eyes driven by a constant-velocity optokinetic stimulus, fluorescence values increased in a ramp-like or exponential fashion; during the subsequent fixation (shaded red), fluorescence levels either remained elevated (e.g. cell 1) or relaxed back toward baseline (e.g. cell 8). Eye position and neuronal activity were reset either by contraversive saccades or contraversive stimulus motion (at right).

3.3.2 Modes of integrator activity during post-saccadic and post-optokinetic fixations

We first determined the persistent firing modes activated during fixations under each behavioral context. To do so, we employed the singular value decomposition (SVD) to analyze, at the population level, the stimulus-triggered average responses of the VPNI population during eye fixations and the preceding stimulus period (Methods). The SVD is a factorization procedure that allowed us to identify, in the state space defined by the activity of all recorded neurons, those directions, or modes, that explain the greatest amount of variation in our data. To identify these directions that captured behaviorally-related variation, we began our analysis by using fluorescence data to perform a rigorous comparison to noise. After significant fluorescence modes were identified, we extracted information about the firing rate modes underlying fixations.

We found that, for each context, fluorescence activity at the population level could be understood by considering the contributions primarily along three modes.

Fluorescence averages were determined over a period that included both the fixation (Figure 3.3a, solid) and a brief period immediately preceding it (Figure 3.3a, dashed). After taking an SVD of the population fluorescence data for each response type separately, we plotted the singular values, or root-sum-squared (rss) amplitudes, of the data projections, or components, along the resulting fluorescence modes (Figure 3.3b, black). To understand how many of these modes were needed to capture the signal in our data, the resulting singular values were then compared to singular values of noise estimated from inter-trial differences[9] (Methods; Figure 3.3b, gray). In each context, only the first three modes had rss amplitudes with confidence intervals that did not overlap with noise.

Next, we applied a deconvolution procedure accounting for the dynamics of calcium buffering in these cells[23,54] to determine, from the significant fluorescence components, three firing modes for each context (Methods). We found that during the fixation period, which was defined by the window between 1 and 7 seconds after termination of the movement command, only two of these modes in each context were required to explain greater than 95% of the mean and variance of the firing rate signal (Supplementary Figure 1). Firing along the third mode was primarily restricted to the period preceding the fixation. During fixations, the primary saccadic and optokinetic firing modes, S1 and K1, were associated with slow relaxation times (Figure 3.3c, magenta traces), and the secondary firing modes, S2 and K2, with faster

dynamics (Figure 3.3c, yellow traces). A linear combination of the contributions along the first two persistent firing modes accurately reflected the trends underlying the relaxation in firing rate during fixations (Figure 3.3d, thick lines; Supplementary Figure 2). We also found that the identified modes did not vary with stimulus amplitude, and that changes in fixation position were represented primarily by changes in the amplitude of firing along the primary modes S1 and K1 (Supplementary Figure 3). Lastly, we confirmed using simulated data with known firing rates the accuracy of this analysis procedure in providing a reliable estimate of the mode structure underlying persistent activity (Supplementary Figure 4). In summary, these results suggest that persistent activity for each context is supported predominantly by two firing rate modes, one with faster and one with slower relaxation times, and that eye position is encoded by the amplitude of firing along the slow mode.

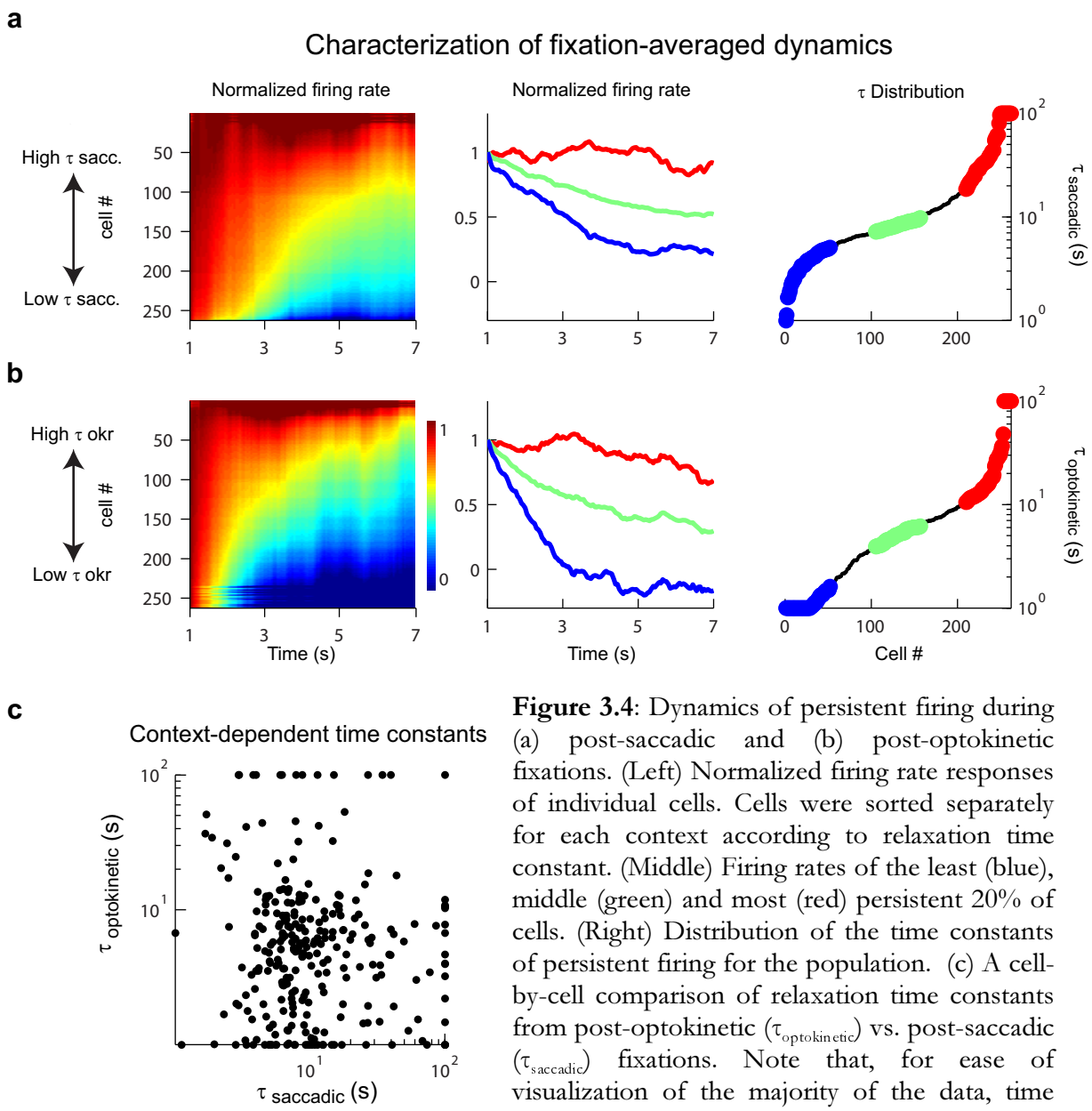


Figure 3.4: Dynamics of persistent firing during (a) post-saccadic and (b) post-optokinetic fixations. (Left) Normalized firing rate responses of individual cells. Cells were sorted separately for each context according to relaxation time constant. (Middle) Firing rates of the least (blue), middle (green) and most (red) persistent 20% of cells. (Right) Distribution of the time constants of persistent firing for the population. (c) A cell-by-cell comparison of relaxation time constants from post-optokinetic ($\tau_{optokinetic}$) vs. post-saccadic ($\tau_{saccadic}$) fixations. Note that, for ease of visualization of the majority of the data, time constants outside the range 1-100 seconds were plotted at the closest corresponding edge of this plot.

3.3.3 Heterogeneity in persistent firing

Next, to see if there were differences between behavioral contexts in the way that the VPNI stored a memory of desired eye position, we examined the pattern of single-neuron firing rates across the population following saccadic versus following optokinetic eye movements. This was done using a traditional measure of persistent firing, the persistence time constant τ, which quantifies the relaxation of activity back to its baseline level after stimulation.

We observed a wide range of time constant values for both behavioral contexts. Heterogeneity at the population level was first visualized for each context separately by using a heat map to display the activities of all neurons (Figure 3.4a and 4b, left), and second by averaging the activity of subsets of neurons (Figure 3.4a and 4b, middle). While the firing rates of some neurons were maintained at approximately the same elevated level throughout the fixation window (top of each panel), the firing rates of other neurons showed moderate relaxation back to baseline (middle), and the firing rates of yet others showed a fast relaxation to baseline (bottom). During post-saccadic fixations (6.9 fixations per cell on average), the mean $\tau_{sacc.}$ was 16.9 seconds, with the lowest, middle, and highest 20% of data exhibiting time constants of 3.6, 8.3 and 52.9 seconds respectively (Figure 3.4a, right). During post-optokinetic fixations (5.6 fixations per cell on average), the mean τ_{okr} was 9.9 seconds, with lowest, middle, and highest 20% values of 1.1, 5.1 and 32.4 seconds (Figure 3.4b, right). Overall, these

data show that at a population-average level the persistent activity for both contexts was similar, suggesting that the saccadic and optokinetic inputs utilized the integrator memory network in much the same way.

However, when we next directly compared the values of the time constants between the two fixation conditions on a cell-by-cell basis, there was a significant surprise: the measures of persistent activity for the two behavioral contexts were *uncorrelated* (Figure 3.4c; Spearman rank correlation, or src=-0.05; p=0.40). We tested whether this lack of correlation was simply due to noise by assessing correlation in the data within a context. We found that time constants between randomly-selected sets of trials from the same context were highly correlated for both post-saccadic (src=0.66; p=0.00) and post-optokinetic (src=0.67; $p=4.7x10^{-35}$) fixations. Thus, although at the population-average level persistent activity during post-saccadic and post-optokinetic fixations appeared quite similar, there must have been *different patterns of persistent firing* across the network for the two contexts.

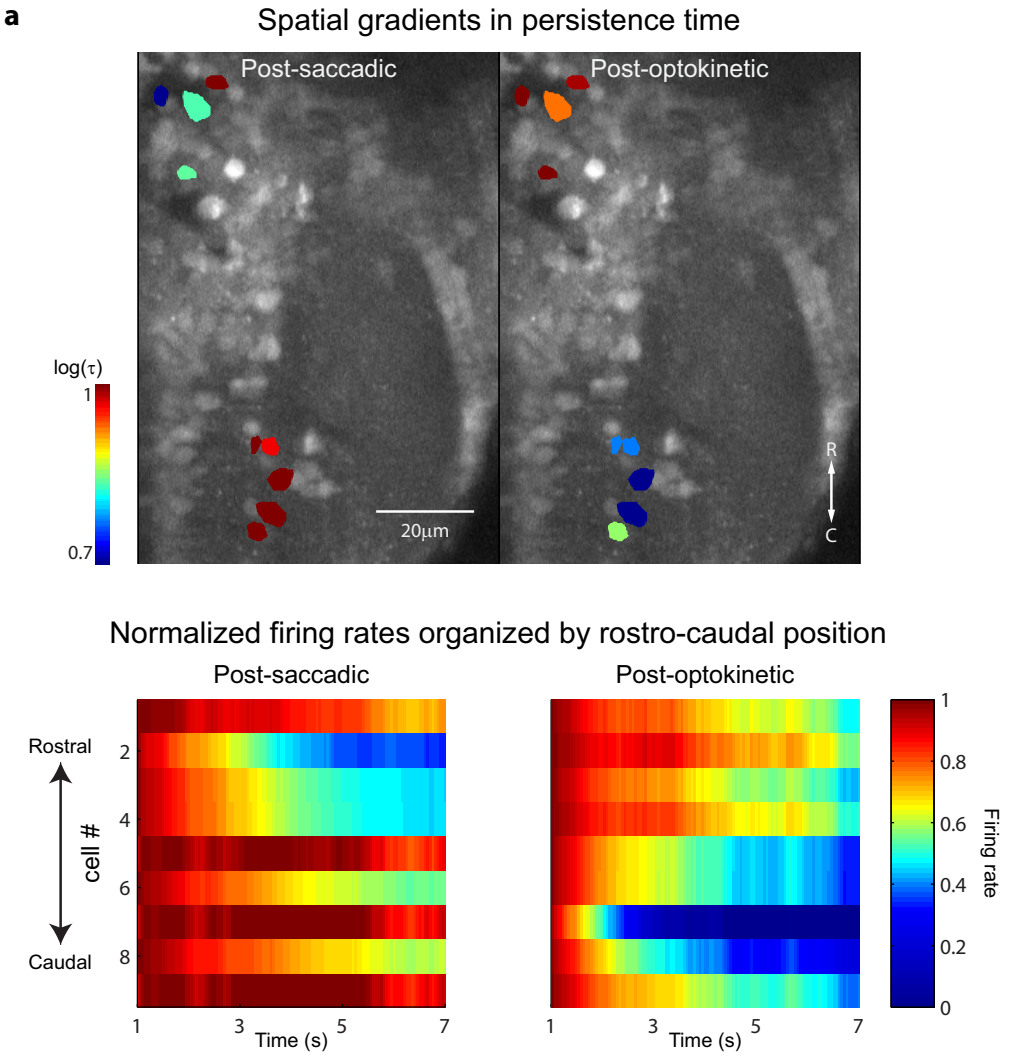


Figure 3.5: Context-dependent reversal in spatial gradients of persistent firing. (a) (Top) Average image at one plane during post-saccadic (left) and post-optokinetic (right) fixations, with VPNI cells color-coded according to $\log_{10}(\tau)$. Note that cells with time constants outside the indicated range were coded with the color of the closest limit. (Bottom) For each context, normalized firing rates for the cells highlighted in the imaging plane above. Cells are arranged vertically according to rostro-caudal position. (b) For the entire population, gradients in persistent firing rate along the rostro-caudal direction assessed by measuring pairwise differences in time constant, normalized by the pairwise mean time constant.

3.3.1 Spatial gradients of persistent firing

Next we sought to determine if the differences between the two fixation conditions in the time constants of persistent firing could be related to different spatial patterns of firing across the integrator network. To do this, we jointly measured the location and time constant of each cell and computed spatial gradients of the measured time constants.

We found an especially striking difference along the rostro-caudal axis: while during post-saccadic fixations the most persistent cells were those located caudally, during post-optokinetic fixations the most persistent cells were instead located rostrally; in other words, there was a significant *reversal* in the spatial gradient of persistent firing along this axis (Figure 3.5a, Supplementary Figure 5). Spatial gradients along the rostro-caudal, medio-lateral, and dorso-ventral axes were first quantified by calculating pairwise differences in time constants (8,890 pairs across 10 fish), normalized by a pair's average time constant (Δ), and comparing to the pairwise distance between the cells. We found that for every 100 micrometers of displacement in the caudal direction (Figure 3.5b, left), in the post-saccadic context persistence values rose by an average of 7.2% (src=0.13, $p=7.1\text{x}10^{-19}$), but in the post-optokinetic context they fell by an average of 7.5% (src=-0.12, $p=1.4\text{x}10^{-15}$). We found that the gradients along the dorso-ventral and medio-lateral axes also differed across the two behavioral contexts, but less notably so (Supplementary Figure 6). In summary, these

data show that there exist differing spatial patterns of firing across the network for the two behavioral contexts, with a significant reversal of gradients along the rostro-caudal direction.

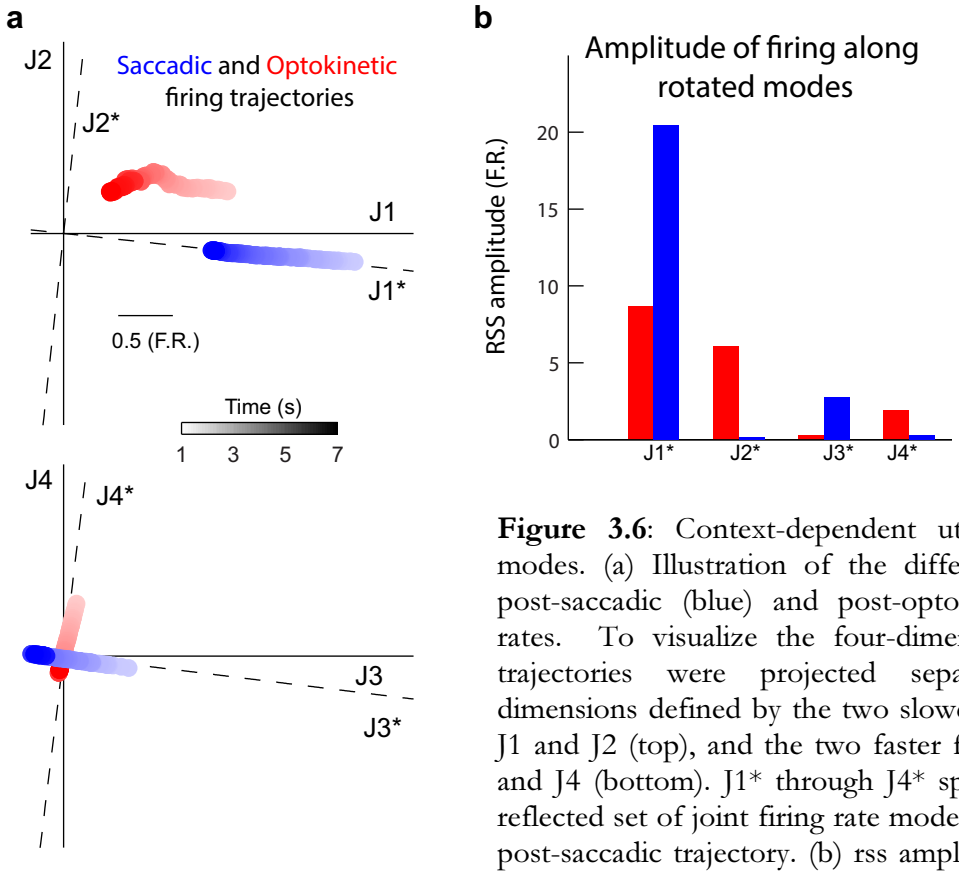


Figure 3.6: Context-dependent utilization of VPNI modes. (a) Illustration of the different trajectories of post-saccadic (blue) and post-optokinetic (red) firing rates. To visualize the four-dimensional trajectories, trajectories were projected separately along the dimensions defined by the two slower firing rate modes J1 and J2 (top), and the two faster firing rate modes J3 and J4 (bottom). J1* through J4* specify a rotated and reflected set of joint firing rate modes that align with the post-saccadic trajectory. (b) rss amplitudes of firing rate projections along modes J1* through J4*.

3.3.2 Degree of overlap between post-saccadic and post-optokinetic modes

The differences in the spatial patterns of activity described above suggest that the two modes active during post-saccadic fixations were not the same as the two modes active during post-optokinetic fixations. In this section we extend our SVD analysis to quantify the degree to which the modes differed.

First, to find a new set of firing rate modes that could explain memory-related activity under both contexts, we performed an SVD on the firing rate signal jointly captured by the modes S1, S2, K1, and K2 during the fixation period (Methods). This identified a set of four orthogonal modes, labeled J1-J4, that were together required to account for at least 95% of both the mean and variance of the firing rate signal. The firing rate components associated with modes J1 and J2 contained information about both the time-averaged level of activity and the slow relaxation of activity in each context, whereas those associated with modes J3 and J4 captured the relatively fast dynamics (Supplementary Figure 7).

Next, to investigate the degree of separation between the fixation signals in the two contexts, we graphically represented their time-varying firing rate trajectories during the fixation period within the new space defined by modes J1 to J4 (Figure 3.6a). Network activity during post-saccadic (blue) and post-optokinetic (red) fixations were plotted in the J1-J2 (top) and J3-J4 (bottom) planes. In the J1-J2 plane, the post-saccadic signal was composed primarily of a component along the J1 direction, while

the post-optokinetic signal received contributions from both the J1 and J2 modes. An even more striking difference was evident in the J3-J4 plane, with the post-saccadic signal following a trajectory nearly perpendicular to the post-optokinetic trajectory. Thus, there was a clear separation of signals in the two contexts, despite the fact that in each case the eye position being represented was highly similar.

We next quantified this separation between the post-saccadic and post-optokinetic signals. To facilitate this comparison, the modes J1-J4 were rotated and reflected so that the post-saccadic trajectory was aligned with the new modes J1* and J3* (Methods; Figure 3.6b, dashed axes). We then calculated the rss amplitudes of the two signals along the new set of modes. As shown in Figure 3.6b, while the post-saccadic signal was almost entirely contained in modes J1* and J3*, and virtually absent from modes J2* and J4*, the post-optokinetic signal was distributed across three modes, J1*, J2*, and J4*. This segregation of post-saccadic and post-optokinetic signals into spatially distinct modes of the network demonstrates that the VPNI represents eye position signals differently depending upon the behavioral context.

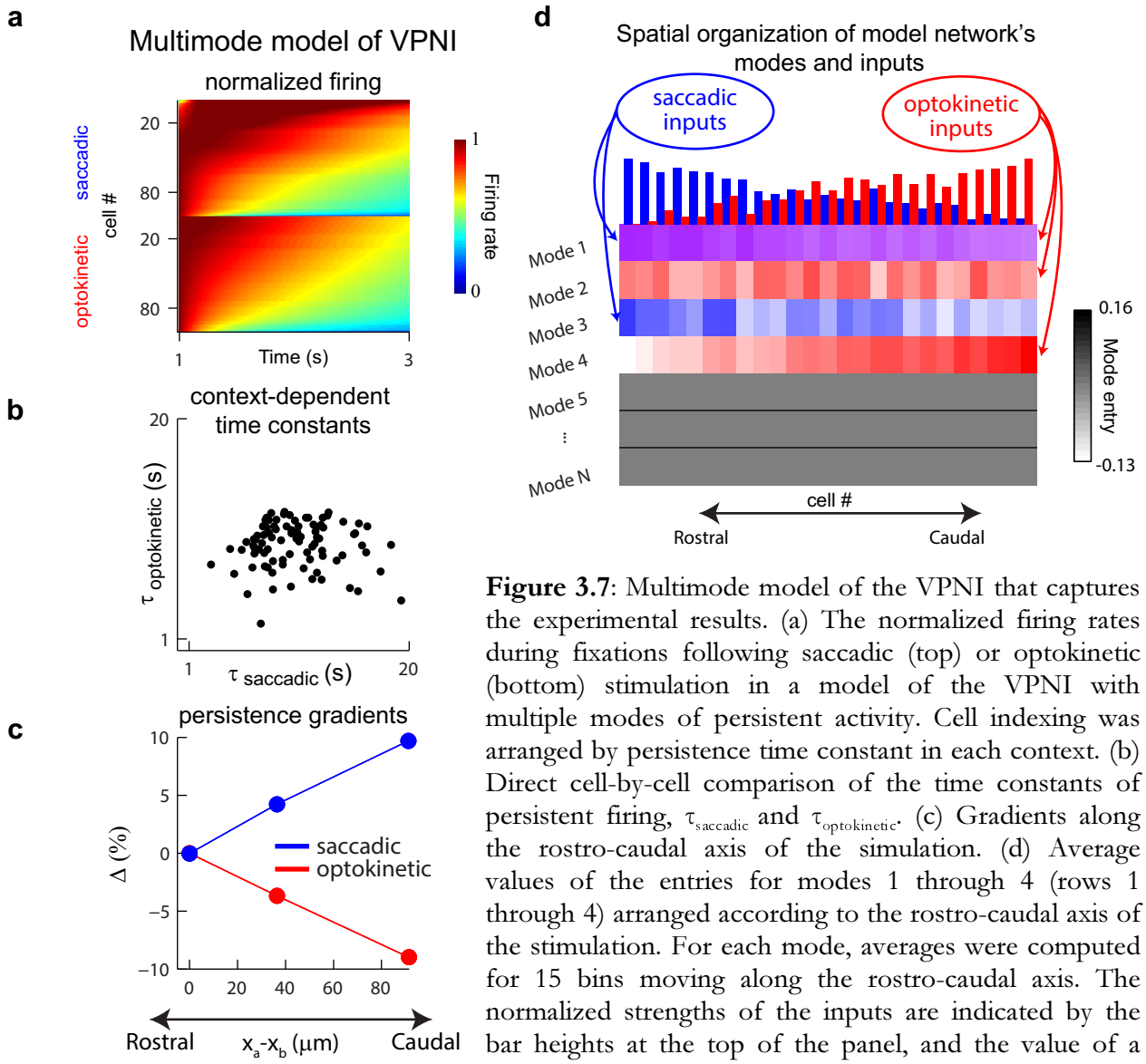


Figure 3.7: Multimode model of the VPNI that captures the experimental results. (a) The normalized firing rates during fixations following saccadic (top) or optokinetic (bottom) stimulation in a model of the VPNI with multiple modes of persistent activity. Cell indexing was arranged by persistence time constant in each context. (b) Direct cell-by-cell comparison of the time constants of persistent firing, $\tau_{saccadic}$ and $\tau_{optokinetic}$. (c) Gradients along the rostro-caudal axis of the simulation. (d) Average values of the entries for modes 1 through 4 (rows 1 through 4) arranged according to the rostro-caudal axis of the stimulation. For each mode, averages were computed for 15 bins moving along the rostro-caudal axis. The normalized strengths of the inputs are indicated by the bar heights at the top of the panel, and the value of a mode entry is indicated by the color intensity.

3.3.3 Potential mechanism of context-dependent persistent activity.

We next explored if and how the current framework for understanding persistent activity in the VPNI memory network could be modified to incorporate our results on context-dependent firing patterns. Previous results from intracellular recordings[25], multi-unit recordings[11,47], and pharmacological perturbations[27] have

suggested that persistent activity and temporal integration in the VPNI is critically dependent on positive feedback generated through local recurrent excitatory connections[8,11,27,48]. Hence, we expect that the patterns of persistent activity observed across the population are intimately linked to the patterns of synaptic feedback in the circuit[8,15]. In the following, we show how the structure of the recurrent synaptic interactions can be arranged to support persistent firing modes that match those found experimentally. We also show that a simple way to get context-dependent access to these modes similar to that seen in our experiments is through spatial pattering of the different inputs: by having stronger saccadic inputs at rostral locations and stronger optokinetic inputs at caudal locations, we find that the multi-mode VPNI model is able to generate heterogeneity in dynamics and context-specific gradients in activity consistent with our experiments.

The VPNI model we constructed was focused on the generation of persistent firing at eccentric eye positions such as those exhibited in the data. The model consisted of 100 recurrently connected neurons, in line with previous estimates of the VPNI population size[24]. We specified the recurrent connection weights so that four modes were present (Methods), with modes 1 and 2 associated with relatively slow dynamics comparable to the above modes J1* and J2*, and modes 3 and 4 with relatively fast dynamics comparable to modes J3* and J4*. In addition, we used the spatial gradients of modes J1*-J4* (Supplementary Figure 8) to specify a rostro-caudal

axis for the model (Methods). Saccadic inputs in the model consisted of 200 millisecond long bursts, and optokinetic inputs consisted of 7-second long steps of constant stimulation, consistent with the experiments (Figure 3.2c). In addition, we chose the input strengths so that the simulated responses approximated the response amplitude distribution in the data (Figure 3.6b).

Context-dependent stimulation of this multimodal model of the VPNI network was able to generate the heterogeneity and gradients of persistence time constant observed experimentally. Context-dependent stimulation was produced by activating modes 1 and 3 with saccadic inputs, and modes 1, 2 and 4 with optokinetic inputs. At the population level, activity for both contexts was composed of cells with long, intermediate, and short persistence time scales (Figure 3.7a). As in the experiments, the cell-by-cell comparison of time constants exhibited poor correlation (Figure 3.7b). The spatial gradients of persistent firing along the rostro-caudal axis exhibited similar trends and a similar reversal between contexts (Figure 3.7c). We found that persistence gradients arose due to the spatial organization of the network's modes (Figure 3.7d), coupled with differential activation of these modes by saccadic and optokinetic inputs that were oppositely graded along the rostro-caudal axis (Figure 3.7d; top bars). In contrast, stimulation of this multimodal network with context-independent inputs failed to replicate all of the experimental findings (Supplementary Figure 9).

We summarize the findings in this work as follows. Our experimental results suggest that the VPNI memory network exhibits multiple modes of persistent activity that can be accessed in a context-specific manner. The modeling results supplement these findings by showing how the existing framework for understanding persistent firing in the VPNI can be extended to incorporate the new data, predicting that differential targeting of VPNI modes is achieved via anatomical gradients in input strength. Together, these results suggest a general strategy for the simultaneous maintenance of stimulus value and stimulus context during short-term memory behaviors.

3.4 Discussion

We have determined a neural basis for context-dependent processing in a short-term memory network. Like other memory systems, the oculomotor integrator is responsible for processing inputs from multiple brain areas. We addressed how this is accomplished by optically imaging the dynamics of large populations of integrator neurons during related behavioral contexts. We found that activity in the integrator during fixations following either spontaneous saccades or optokinetic ramps was explained by one quickly relaxing and one slowly relaxing mode of persistent firing. However, a comparison of the modes activated during saccadic versus optokinetic behaviors revealed them to be quite different, and this difference was manifest most prominently as a reversal in a spatial gradient of the relaxation time constant. Taken

together, these results provide an example of how context-dependent processing in a short-term memory network can be accomplished through the utilization of different modes of persistent neural activity.

This work has shown that both the overall amplitude of persistent firing and the pattern of relative firing rates can be used to represent information. Graded persistent activity has historically been studied at the single-neuron level, where the main focus has been on characterizing the relationship between a single memory variable and the amplitude of firing. The presence of this relationship has been seen as evidence that persistent firing is used for storing the value of a stimulus, or *what* (stimulus) information. Our work shows that, in addition to storing stimulus value, the VPNI network retains information about *how* a memory had been activated by using distinct patterns of persistent activity to represent different contexts. We suggest that this strategy of using multiple patterns, or modes, could be important for memory systems in general, where, in addition to retaining stimulus value, a network may need to simultaneously encode *how* the stored information is to be used[55,56,57], *when* a behavior is to be performed[9], and *where* a stimulus was located[58]. On a broader level, utilization of multiple modes to generate different input-output relationships also could be applicable to non-memory settings requiring multi-sensory computations[59,60,61].

There are several reasons why the VPNI might use multiple modes for storing information. One potential reason may lie in a division of labor at the level of the downstream motoneurons and muscles of the oculomotor system. Studies in the tadpole[62,63] and primate[64,65] suggest that quickly fatiguing twitch fibers receive input from motoneurons with large somatic and axonal diameters, while fatigue-resistant slow non–twitch fibers receive inputs from smaller motoneurons. Because saccadic and optokinetic eye movements are performed at vastly different speeds, optimal control of eye movements might require the recruitment of a larger proportion of twitch fibers during saccades and non-twitch fibers during OKR. If so, the oculomotor plant could have different dynamical properties after the two stimuli[66], which would require differential neural control for seemingly identical fixations. Considering the large and small motoneuron populations as separate targets of the VPNI, differential control could be established by the activation of different subsets of persistent firing modes. Likewise, given the wide range of brain areas that receive efference copy signals from the VPNI, including the cerebellum, thalamus[42], and posterior parietal cortex[43], it is reasonable to speculate that different input-output relationships in the VPNI could be used to alter different aspects of a motor plan[67,68].

VPNI multimodality also might be useful for the control of learning. The time constant of oculomotor integration at the behavioral level can be modified by visual feedback[69], a process mediated by the cerebellum (O. Debowy and R. Baker,

unpublished observations) that involves changes in the firing of VPNI neurons[70]. The presence of multiple modes in the integrator could enable the cerebellum to control the behavioral time constant of integration purely by modifying the weights of input connections to the VPNI, adjusting the relative strengths of saccadic or optokinetic stimulation in a mode specific manner. Such an approach would be easier to implement than the traditionally proposed mechanism of retuning feedback within the VPNI network[71,72] as it would only require the modification of N connections from external inputs as opposed to N^2 recurrent connections.

The mode analysis presented here revealed a clear separation between the fixation signals following saccadic input and those following an optokinetic stimulus ramp. We did observe significant overlap along one mode, J1*. In principle, this overlap could be related to the fact that, for both contexts, VPNI fixation activity was measured while the stationary optokinetic stripe pattern was present. This was done to ensure that post-saccadic and post-optokinetic signals were monitored under the same conditions. This did, however, generate a source for common signal into the VPNI during fixations, retinal slip. Although we did not attempt to quantify how much of the overlap along mode J1* was related to this common signal, given the similarities between dark[50] and light[11] conditions in fixation dynamics both before and after VPNI inactivation, it is likely that the influence of this signal upon integrator firing during fixations, and therefore any contribution to the overlap, was quite small.

We also explored how the current framework for understanding the generation of persistent activity in the VPNI might be extended to incorporate our experimental results. Previous experimental and theoretical work investigating persistent firing in the VPNI has supported a prominent role for local recurrent excitatory interactions. Intracellular voltage measurements have shown that eye movements to positions at which an integrator neuron would exhibit high rates of persistent firing were accompanied by increases in the mean and variance of membrane potential, suggesting a sustained increase in the rate of synaptic inputs[25]. Electrical recordings of neighboring pairs of integrator neurons revealed that spike trains were positively correlated with lags shorter than 10 milliseconds, consistent with excitatory synaptic coupling[26]. Furthermore, disrupting this coupling through localized pharmacological inactivations in one portion of the network resulted in a decrease in the time constant of persistent firing for neighboring cells in non-inactivated portions, suggesting a causal role for recurrent excitation in generating persistent activity[27]. These data support an attractor network formalism that has a long history in the VPNI[15,27,44,45] and other short-term memory settings[20]. In this framework, patterns of persistent activity across the network are established by the patterns of synaptic feedback interactions between constitutive neurons, with interactions specified by both synaptic strengths, and static or dynamic nonlinearities such as short-term plasticity[73,74] and intrinsic cellular response properties[8,27,45,75]. Our proposed model expands the classic

"line attractor" formalism of the VPNI[15], which posited that a single mode of activity was able to explain persistent firing during fixations, to a "multimodal attractor" with at least four modes of activity that is able to replicate our experimental findings on heterogeneity of time constants and context-dependent spatial patterns in persistent firing. In the model, we demonstrated how these four modes can be generated easily through specification of four distinct patterns of connectivity within the network. We also showed how context-dependent processing in this system can be initiated simply through opposing gradients in the transient saccadic and optokinetic inputs.

It is further possible in this attractor network framework to consider alternatives for generating and accessing different persistent firing modes. As noted above, and as explored more fully in our recent, biophysically-detailed model of the VPNI system[8], the interactions between integrator neurons may be governed by more than the patterns of recurrent synaptic weights. For example, if recurrent inputs were mediated by activation of a voltage-sensitive dendritic current, then the mechanism of persistent firing would be composed of both a "network" component specified by synaptic connection strengths, and a "cellular" component specified by the intrinsic kinetics and distribution of the channel mediating the dendritic current. Thus, although the simplified model of Figure 3.7 is able to capture all of the experimental findings reported here solely through specification of the patterns of recurrent and external inputs, future work will be needed to investigate if one or more cellular

parameter could be used to mediate feedback in a way that would account for our current results. One interesting possibility to pursue is the idea that the two fast modes identified reflect the presence of distinct cellular processes[27], one activated by saccadic inputs, and the other by optokinetic inputs. If this were the case, since firing rate dynamics were established by a linear combination of short and long modes (Supplementary Figure 2), gradients in persistent firing would be of primarily cellular origin as it was the relative contributions of the short modes that varied the most (Figure 3.7d, Supplementary Figure 8). However, such gradients in intrinsic properties cannot fully explain our current results, as they do not explain the origin of two modes with slow dynamics. Another possibility is inspired by work in the rhythm-generating circuits of the stomatogastric ganglion showing how large numbers of different neuromodulators may together control many aspects of circuit dynamics[76,77], and studies suggesting cholinergic modulation of CAN[78,79] or HCN[80] currents is important in persistent firing. In principle, neuromodulators associated with the separate inputs could alter cellular and network properties together to drive transitions in the system between two sets of modes. However, saccadic and optokinetic inputs often occur simultaneously in this system, as optokinetic ramps were typically accompanied by saccades, and during spontaneous saccadic behavior the vestibular neurons providing optokinetic signals fire at high tonic rates[81]. Therefore, differences in persistent activity between the two contexts would have to depend on subtle

differences in the concentrations of the two neuromodulators, and thus the magnitude of inputs. This would be difficult to align with our finding that the patterning and dynamics of persistent activity was independent of input amplitude (Supplementary Figure 3). Furthermore, transitions between the different behavioral contexts and associated persistent activity modes generally occurred within seconds, a time scale that is faster than those typically associated with neuromodulatory effects[82,83]. Therefore, we suggest that the model proposed in Figure 3.7 based on patterning of connectivity represents the most parsimonious hypothesis consistent with our current and previous[25–27] experiments in this system.

In conclusion, we have addressed the question of how temporal integrators generate short-term memory under different behavioral contexts. We have shown experimentally that the VPNI memory circuit exhibits at least four modes of persistent firing, and that these modes can be accessed in a context-dependent manner. Through this dynamical organization, the VPNI is able to store the value of a memory in the overall amplitude of firing, and represent the manner by which the memory was activated and will be used in the spatial patterns of firing across the circuit. We also have shown how the current framework for understanding integration in the VPNI can be modified to incorporate our new findings, defining patterns of recurrent synaptic connectivity that allow the network to support multiple modes of persistent firing. Interestingly, similar modifications could easily be made to provide

specificity for additional contexts, offering a potentially scalable solution for the challenge of context-dependent information processing in short-term memory networks.

3.5 Methods

3.5.1 Zebrafish preparation and dye-loading

All experimental procedures were approved by Weill Cornell Medical College's Institutional Animal Care and Use Committees. Nacre[50] mutant zebrafish larvae, 5-9 days of age, were used in all experiments. Larvae were anesthetized in 100 or 200 mg/L tricaine-methanesulfonate (MS222, VWR TCT0941-025G) for about 1 minute and then quickly mounted dorsal side up with droplets of 1.7% low-temperature agarose (Sigma A0701-100G) in the lid of a 35mm petri dish containing a bed of hardened 1% agarose (Invitrogen 15510-027). Animals were then covered in 50 mg/L MS222. Bulk-loading[84,85] of the calcium sensitive dye Oregon Green 488 BAPTA-1 AM (5 mM, in DMSO with 5% pluronic) was performed within 20 minutes of embedding. The dye (Invitrogen, O-6807) was previously dissolved to 10 mM in DMSO and stored in aliquots of 0.25 μL, at -20°C. The pluronic surfactant (F-127, Invitrogen, P6878) was prepared at a 10% concentration in DMSO no more than a few hours prior to injections. Immediately before injection a dye aliquot was removed from the freezer, hand-warmed, and combined 1:1 with the pluronic solution. After mixing, the dye/pluronic solution was immediately loaded into a glass capillary (WPI,

1B100F-4WPI) which was prepared using a vertical puller (Sutter P-30), with the tip diameter broken to ~5 μM. Under 10x magnification (Leica S6E, Leica KL1500 illumination), the capillary was then inserted through the dorsal skin surface over the lateral edge of the right side of the hindbrain just caudal to the cerebellum. The tip was then advanced diagonally at a ~30 degree decline relative to the dorsal surface toward the lateral edge of the left side of rhombomere-7 or -8, a region largely composed of neuropil containing the dendrites of integrator neurons (unpublished observations). The dye solution was then pressure injected using 150-250 ms pulses, with the pressure adjusted (2-10 PSI) so that refractive-index-mediated changes from the plume of dye were visible. Typically 10 pulses were given within the course of about 1 minute. Immediately following injection, fish were removed from agarose for several hours in order to recover from anesthesia.

3.5.2 Two-photon calcium imaging during fixation behavior

Two-photon imaging and eye-tracking were performed using custom-designed instrumentation; most features of this setup have been previously described[50]. Briefly, laser power control was achieved using an electro-optical modulator (Conoptics 350-50UV, driven using Conoptics 302 RM amplifier). Excitation light (790 nm) was provided through a 40x objective to the caudal hindbrain, while illumination of the eyes was achieved using an infrared LED (850 nm) aimed at the back of the same objective. Eye movements were monitored using a sub-stage CMOS camera (Allied

Vision Technologies, Guppy FireWire Camera) at a sampling rate of approximately 13 Hz. Eye position calculations were performed using a custom Matlab program. Because the movements of the left and right eyes were highly coordinated, as observed previously[50], we only monitored the movements of the eye with the best signal-to-noise ratio position calculations; typically this was the left eye, ipsilateral to the large majority of VPNI neurons loaded with the calcium sensor.

An optokinetic stimulus of vertical stripes was provided via a small projector (3M, MPro110). The projected light was passed through a bandpass (Thorlabs FGB25, 315-445 nm) and short-pass filter (Thorlabs FES0450, 450 nm) so that it would not interfere with fluorescence collection or eye tracking. The resulting output of black stripes on a blue background was displayed on a screen placed roughly 2 cm in front of the fish; the screen consisted of a 10 cm x 1 cm piece of diffusion film (Novatron 4008 Diffusion Sheets) anchored to the microscope stage using wax (Wiki Stix). A projected stripe measured roughly 1 cm in width on the screen, corresponding to approximately 30 degrees of the animal's field of view. Custom Matlab software was also used for stimulus control. The microscope and above instrumentation were located within a light-tight enclosure.

Fluorescence imaging was performed in the horizontal plane bilaterally within rhombomeres 7 and 8 over an area measuring 184 μm on a side. For each animal, we

imaged a stack of 3-7 planes separated by 10-15 μm. Each plane was monitored for 10 sequential minutes at approximately 1 Hz (512 by 512 pixels, 2 ms/line).

The behavioral paradigm used was designed to test for differences in the patterns of VPNI activity during fixations under two different contexts. First, spontaneous saccades and post-saccadic fixations were generated by the zebrafish while the optokinetic stimulus was stationary (5 minutes). Next, the optokinetic stripes were moved to either the fish's right or left side at a constant velocity within the range of 10-20 degrees/sec, with only a single velocity used for each fish. Post-optokinetic fixations were produced by manually setting the stimulus velocity to zero after the eye had reached roughly the same eccentric position as achieved during spontaneous saccades. Note that since the larval zebrafish does not generate optokinetic after-nystagmus[86], no velocity-storage signal would be provided to the VPNI during fixation; hence, the two fixation conditions should be identical (Supplementary Figure 10). Following a fixation period of about 10 seconds, the stimulus motion was resumed and the process was repeated a number of times in both directions (5 minutes). Typically, the time between acquisitions of saccadic and optokinetic responses was about 10 seconds. Given previous findings on the organization of recruitment thresholds in the fish VPNI[24], it is expected that all neurons on a particular side of the brain were brought above threshold as the eyes moved to the ipsilateral hemifield.

Integrator cell fluorescence responses and selection criterion

Putative integrator neurons were identified with a modified version of a previously described method[50]. Briefly, for every pixel in an image, we first determined the correlation over the 5 minute acquisition period with each of the following variables: eye position, eye velocity, and, for optokinetic behavior, stimulus position and stimulus velocity. Regions of Interest (ROI) were then drawn by a user interface around somata where the majority of pixels were correlated with at least one of these variables. The fluorescence time series for a cell was computed by taking the average fluorescence over its ROI for each frame. Traces were converted from fluorescence intensity to fractional change in fluorescence by taking the difference of each trace from the average fluorescence, and then dividing by the average. Finally, because the fluorescence traces occasionally drifted on the minutes timescale, a baseline correction was employed. For each cell, the full activity time series during saccadic and optokinetic eye movements were separately fit with a linear combination of eye position and eye velocity convolved with the calcium impulse response function. This model has been shown to capture the majority of the variance in the fluorescence activity of VPNI neurons[50]. To account for slow fluctuations in the fluorescence unrelated to eye movements, an additional third order polynomial term was added to the model. This polynomial was then subtracted from each fluorescence time series and zero was set to coincide with the average of those data points with

intensity in the 20th percentile. To be included in the subsequent analyses, the R^2 value of these fits during saccadic eye movements had to be greater than 0.25 and cells needed to have a correlation with eye position during saccadic and optokinetic behavior that was greater than 0.2 (Pearson correlation).

Next, we applied a cell-selection step based on the number of fixations available for calculating stimulus-triggered fluorescence averages. Fixation criteria were defined relative to the primary or null position, which was calculated as the average of a subset of positions from the left (5%) and right (5%) extremes. For post-saccadic fixations, the eye had to be in the ipsilateral hemifield at positions no greater than 16 degrees ipsilateral to the null. For post-optokinetic fixations, the preceding optokinetic ramp had to follow a contraversive saccade, and the eye had to be at least 3 degrees beyond the primary position in the ipsilateral hemifield. These conditions ensured that distributions of post-saccadic and post-optokinetic fixation positions did not differ significantly. Cells were included in subsequent analyses if there were three or more such qualifying fixations for each behavior.

To facilitate averaging across trials, the fluorescence traces were first interpolated so that their new time points matched those of the eye traces[11]. Segments of interpolated data from qualifying trials during the fixation and pre-fixation periods were then excised and averaged together to generate saccade- or optokinetic-triggered fluorescence averages. For optokinetic eye movements, fluorescence traces were

analyzed during the 11 second window beginning 4 seconds prior to the end of the stimulus. Saccadic fluorescence traces were analyzed during the 8 second window starting 1 second before the end of the saccade. The analysis of firing rates was performed during the fixation period specified as the 6 second window beginning one second after the end of the stimulus. We chose this start time for fixation analysis because the interpolation procedure led to smoothing that could allow signals at the end of stimulation to inappropriately "spill over" into times after the stimulus.

Three additional selection steps were applied. First, in association with measurements of the calcium impulse response function (see "Assessment of persistent firing"), fluorescence following contraversive saccades had to be well fit by an exponential function ($R^2 > 0.3$). Second, saccade-triggered average fluorescence during the 4-second window centered on the occurrence of the saccade had to be well correlated with eye position or velocity convolved with the calcium impulse response function (Pearson correlation coefficient>0.6). Third, firing rates normalized by their average value at the start of the fixation (see below) could not have average values that were larger than 1.3 or smaller than -0.3, thereby eliminating cells whose activity was excessively noisy. The SVD procedure below was repeated following elimination of such cells from the population.

3.5.3 Determination of mode structure through the Singular Value Decomposition

The singular value decomposition (SVD) was used to identify the state-space trajectories traversed by the VPNI during post-saccadic and post-optokinetic fixations. Here we briefly summarize our methodological steps before providing complete details. As a first step, we performed an SVD on the fluorescence data encompassed within both the fixation period as well as the stimulus period immediately preceding it. We then determined which modes contained significant stimulus and fixation related fluorescence signals by comparing to an SVD of the noise estimated in our data. Next, we deconvolved the fluorescence signal using a kernel representing the dynamics of calcium buffering in these cells[50,87], thereby obtaining an estimate of the firing rate changes associated with the stimulus and fixation periods. We then identified those firing rate modes contributing significantly to persistent activity. Note that fluorescence responses, rather than firing rate estimates, were used at the initial stages of this analysis because deconvolving fluorescence results in amplification of high frequency noise, making assessment of the contribution of noise to the dimensionality of the data more difficult. Results from performing an SVD on simulated data revealed that the dimensions identified from this approach should be nearly identical to those present in the underlying firing rates (Supplementary Figure 4).

For both contexts, we first performed an SVD on the responses during that context alone. To do this we constructed a $T \times N$ matrix of fixation-averaged fluorescence responses, $\mathbf{F}$, where columns correspond to cell number (total of N cells) and rows correspond to time points (total of T time points). For saccadic and optokinetic fluorescence responses the respective time ranges were from -1 to 7 seconds and -4 to 7 seconds. These matrices incorporated data from all cells over all fish. The SVD was computed such that

$$\mathbf{F} = \mathbf{USV}', \qquad 3.1$$

where $\mathbf{V}$ is an $N \times N$ matrix whose columns form an orthonormal basis of right singular vectors or modes (total of N modes), $\mathbf{U}$ is a $T \times T$ matrix of left singular vectors, and $\mathbf{S}$ is a $T \times N$ matrix whose diagonal entries s_n are the singular values of $\mathbf{F}$.

The projections of the activity along the fluorescence modes, referred to as temporal fluorescence components, were computed by multiplying the matrix of responses $\mathbf{F}$ by the matrix of modes $\mathbf{V}$ such that

$$C_{t,n} = \sum_{i=1}^{N} F_{t,i} V_{i,n} = U_{t,n} s_n, \qquad 3.2$$

where $F_{t,i}$ is the average fluorescence value at time t of cell i, $V_{i,n}$ is the entry in mode n specific for cell i, $U_{t,n}$ is the t^{th} time point of the n^{th} left singular vector, and $\mathbf{C}$ is a $T \times N$ matrix whose n^{th} column is the component along the n^{th} mode.

To measure the strength of the projection along a given mode we computed the root-sum-square (rss) amplitude, A, of the corresponding component. Therefore, the strength of the projection along mode n was given by

$$\begin{aligned} A_n &= \sqrt{\sum_{t=1}^{T} C_{t,n}^{\ 2}} \\ &= \sqrt{s_n^2 \sum_{t=1}^{T} U_{t,n}^2} \\ &= \begin{Bmatrix} s_n, & n \le T \\ 0, & n > T \end{Bmatrix}, \end{aligned} \qquad 3.3$$

where the last step follows from the orthonormality of the matrix **U**. Therefore, the rss amplitude of the n^{th} component is given by the n^{th} singular value.

The uncertainties in singular values were computed using a bootstrap analysis in which 1000 distinct singular value spectra were computed for 1000 distinct T x N bootstrap data sets. Each data set was compiled by randomly selecting the response of a cell from our N cell population and repeating the process anew 1000 times. The distribution of singular values for each mode calculated from these matrices was characterized by measuring the 90% confidence intervals.

Since the fluorescence activity that supplied our SVD analysis consisted of both signal and noise, we wanted to determine which of the identified modes contributed to the signal. To do so, we compared the rss amplitude along each mode to the magnitude expected from noise. Noise was estimated by implementing a methodology put forth by Machens *et al.*[88], as follows:

The fixation response of a cell can be thought of as the sum of a signal of interest, *g*, and a term accounting for noise, *δ*. For a cell *i*, at time point *t*, during the l^{th} of *L* trials, the fluorescence response $f_{t,i}^{l}$ is then

$$f_{t,i}^{l} = g_{t,i} + \delta_{t,i}^{l}. \qquad 3.4$$

The trial-averaged data, $F_{t,i}$, can then be written

$$\begin{aligned} F_{t,i} &= \frac{1}{L}\sum_{l=1}^{L} f_{t,i}^{l} \qquad 3.5 \\ &= g_{t,i} + \bar{\delta}_{t,i} \end{aligned}$$

where $\bar{\delta}_{t,i}$ is the trial-averaged noise.

Ultimately, we are interested only in the state-space trajectory of the signal g, so it is important to establish an estimate of $\bar{\delta}$. This can be done by taking the difference between the responses of two trials *l* and *q* since $f_{t,i}^{l} - f_{t,i}^{q} = \delta_{t,i}^{l} - \delta_{t,i}^{q}$. Since this quantity is the difference of two random variables, it will itself be a random variable with a standard deviation that is a factor of $\sqrt{2}$ greater than that of a single generic trial *p*; in other words $\delta_{t,i}^{l} - \delta_{t,i}^{q} = \sqrt{2}\delta_{t,i}^{p}$. Because the quantity $\bar{\delta}_{t,i}$ is the average of *L* such terms, we have $\bar{\delta}_{t,i} = \frac{\sqrt{L}\delta_{t,i}^{p}}{L}$. Therefore, the average noise trace $\bar{\delta}_{t,i}$ is given by

$$\bar{\delta}_{t,i} = \frac{\left(f_{t,i}^{l} - f_{t,i}^{q} \right)}{\sqrt{2L}}. \qquad 3.6$$

We can then estimate the contribution of the noise by computing the SVD of the matrix $\bar{\delta}$. The largest noise singular value, s_1^{δ}, represents the amplitude of the noise along its largest dimension. Confidence intervals for these noise singular values s_n^{δ} were calculated as above for the data.

Assessment of persistent firing

We next identified the persistent firing rate signal associated with the fixation period. To do so, we first deconvolved the fluorescence signal identified by the above analysis to eliminate the effects of calcium buffering[11] and determine the firing rate signal associated with the stimulus and fixation periods. We then applied a second SVD to the firing rate signal associated with the fixation period alone to identify a set of orthogonal firing rate modes necessary to explain persistent activity.

When the saccadic and optokinetic contexts were considered separately, the comparison between fluorescence data and fluorescence noise over the window that included stimulus and fixation related responses revealed that 3 modes were needed to explain the signal (Figure 3.3b). Together with equations 3.1 and 3.2, this means that the fluorescence signal of the i^{th} cell, $F_{t,i}^{sig}$, was given by

$$F_{t,i}^{sig} = \sum_{n=1}^{3} V_{i,n} C_{t,n} , \qquad 3.7$$

the contribution received from the 3 significant modes.

The firing rate signal of cell *i*, $r_{t,i}^{sig}$, was determined by deconvolving $F_{t,i}^{sig}$ using a kernel termed the calcium impulse response function, or CIRF[31], that describes the low-pass filtering due to calcium buffering that predominantly governs the transformation from underlying firing rate to calcium fluorescence. This CIRF was calculated by fitting to contraversive saccade-triggered fluorescence averages, as described previously[11].

Next, we performed an SVD on these firing rate responses during the fixation period between 1 and 7 seconds after the offset of the movement command. This specified set of orthogonal persistent firing rate modes $\mathbf{V^r}$, together with their firing rate components $\mathbf{C^r}$. Significance of the n^{th} firing rate mode was assessed through computing the fraction of the total variation explained by the n^{th} component, $C_{t,n}^{r}$, relative to both zero firing (γ_n^{zero}) as well as the component's time-averaged value (γ_n^{mean}). These quantities were given by the following two equations:

$$\gamma_n^{\text{zero}} = \frac{\sum_{t=1}^{7} \left(C_{t,n}^{r} \right)^2}{\sum_{k=1}^{3} \sum_{t=1}^{7} \left(C_{t,k}^{r} \right)^2} , \qquad 3.8$$

And

$$\gamma_n^{\text{mean}} = \frac{\sum_{t=1}^{7}\left(C_{t,n}^r - \langle C_{t,n}^r \rangle_t\right)^2}{\sum_{k=1}^{3}\sum_{t=1}^{7}\left(C_{t,k}^r - \langle C_{t,k}^r \rangle_t\right)^2}. \qquad 3.9$$

Because the results presented in Supplementary Figure 1 revealed that fixation-related firing was significant only along the first two modes in each context, the firing rate signal for cell *i* during the fixation period, $r_{t,i}^{fix}$, was given by the equation

$$r_{t,i}^{fix} = \sum_{n=1}^{2} V_{i,n}^r C_{t,n}^r, \qquad 3.10$$

where $V_{i,n}^r$ refers to entry *i* of firing rate mode *n*, or equivalently, the weight of firing rate component *n* for cell *i*.

In Figure 3.3c, we compared the relaxation dynamics towards baseline of the firing rate components associated with the primary and secondary modes of the network activity. This was done by fitting the first component to the function $ae^{-t/\tau}$, and the secondary component to the function $ae^{-t/\tau} + c$. For illustrative purposes only, the two fits were aligned by their baseline value. For Figure 3.4 and Figure 3.5,

persistent firing rates during fixations were then quantified by determining the relaxation time, τ, resulting from fits of $r_{t,i}^{fix}$ to the form $ae^{-t/\tau}$.

3.5.4 Spatial gradients in persistent firing

Gradients in persistent firing were calculated by taking the pairwise difference $\tau_a - \tau_b$ between the relaxation time constants of cell *a* and cell *b*. This difference was then divided by the pair's average time constant to yield the percentage difference, Δ such that

$$\Delta = \frac{\tau_a - \tau_b}{\frac{1}{2}(\tau_a + \tau_b)} \times 100 \qquad 3.11$$

This quantity was plotted against the pairwise distance between the two cells, $x_a - x_b$, along the rostro-caudal, medio-lateral, and dorso-ventral axes. Pairwise measurements were made only for those cells on the same side of the brain in a given fish. Statistics were calculated by pooling all pairs from all fish. Gradient plots were made by sorting data points in order of increasing pairwise distance and then separating into bins containing an equal number of pairs. Error bars were determined as the standard error of the mean.

3.5.5 Overlap between post-saccadic and post-optokinetic modes

To identify a region in state-space that contained the post-saccadic and post-optokinetic signals together we performed a joint SVD on the firing rate signals during post-saccadic and post-optokinetic fixations identified from the analysis of Figure 3.3. First, the post-saccadic and post-optokinetic fixation responses were concatenated to form a $2T$ x N matrix, where T represents the number of time points in the trial-averaged saccadic and optokinetic fixations. An SVD of this joint signal matrix revealed that four modes, labeled J1-J4, were required to explain at least 95% of the fixation-related firing (Supplemental Figure 3.7). To quantify the separation between the firing trajectories during post-saccadic and post-optokinetic fixations within this 4-D space, modes J were then rotated to a basis J^* where two of the modes were aligned closely with the primary and secondary post-saccadic modes identified in Figure 3.3. The sign of mode J4* was reversed to ensure that the sign of the initial post-saccadic firing rate projection was positive. In the resultant basis defined by modes *J1** through *J4**, *J1** was almost perfectly aligned with the primary saccadic mode S1, and *J3** with the secondary mode S2. By then measuring the size of the post-saccadic and post-optokinetic signals' projections along these new modes, we were able to assess the separation between the signals associated with each context (Figure 3.6b).

3.5.6 Model simulations of a multi-modal integrator

To generate a network capable of reproducing the main features of our data, we used a linear model of the form:

$$\tau \frac{dr_i(t)}{dt} = -r_i(t) + \sum_{j=1}^{N} W_{ij} r_j(t) + h_i^{sac} I^{sac}(t) + h_i^{okr} I^{okr}(t) \qquad 3.12$$

where $r_i(t)$ is the firing rate of neuron i, W_{ij} is the connection from neuron j onto neuron i, τ is the cellular time constant, which was taken to be 100 milliseconds, $I^{sac}(t)$ and $I^{okr}(t)$ are the functional forms of the saccadic and optokinetic inputs to the network, and h_i^{sac} and h_i^{okr} are their strengths onto neuron i. The network consisted of 100 neurons, in line with experimental estimates of the number of cells in this system[24].

The connectivity **W** was specified by determining a rank-4 matrix in which two modes supported persistent activity with a ten second time constant and the other two modes supported activity with a one second time constant, consistent with our joint SVD analysis (Figure 3.6). To do so, we first determined an appropriate set of orthonormal modes or eigenvectors, then constructed **W** from these eigenvectors and their corresponding eigenvalues. We began with four arbitrary vectors, three with entries randomly selected from the range 0 to 1, and a fourth where each entry was

given by 1 minus the corresponding entry of the first vector. The Gram-Schmidt operation was then applied to this set to generate four orthonormal eigenvectors, $\vec{E}_1$ to $\vec{E}_4$. Next, the connectivity matrix **W** was formed using the outer product $W_{i,j} = \sum_{k=1}^{4} \lambda_k E_{i,k} E_{j,k}$, where the term $E_{j,k}$ represents the j^{th} entry of the k^{th} eigenvector; the eigenvalues λ_k were selected to be 0.99, 0.99, 0.9 and 0.88, respectively, for eigenvectors 1 through 4, to approximate the component dynamics observed experimentally (Supplementary Figure 7).

Network simulations were run with conditions specified according to the experimental protocols. To generate a saccade, we set $I^{sac}(t)$ to 1 for 200 milliseconds; to generate an optokinetic ramp, we set $I^{okr}(t)$ to 1 for 7 seconds. In both cases, after stimulation, external input was set to zero. To simulate context-dependent activation of different modes in a manner consistent with the data of Figure 3.6, saccadic input strengths were chosen to be $\vec{h}^{sacc} = 4\vec{E}_1 + \vec{E}_3$, while optokinetic strengths were set to $\vec{h}^{okr} = 2.5\vec{E}_1 + 1\vec{E}_2 + 5\vec{E}_4$.

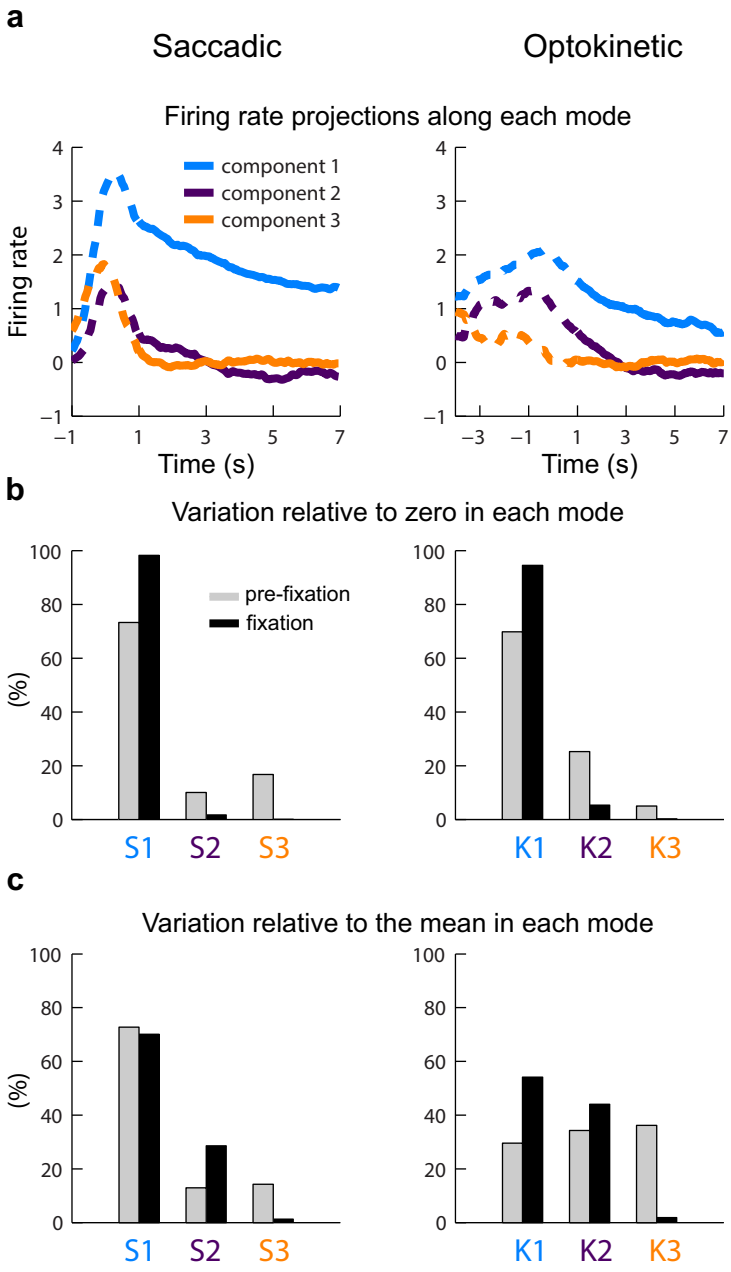
Next, to help firmly connect the simulation results with the experimental findings, we constructed a rostro-caudal axis so that the vectors $\vec{E}_1$ to $\vec{E}_4$ exhibited spatial gradients similar to those of the experimental modes. Although spatial gradients in persistent activity were present along the dorso-ventral and medio-lateral axes (Supplementary Figure 6), because the gradients and differences along the rostro-

caudal axis were most pronounced, for simplicity and without loss of generality, we focused the model's analysis only on this direction. Construction of this spatial axis was achieved by assuming that the modes $\vec{E}_n$ could be approximated by the linear relationship:

$$E_{i,n} \approx a_n X_i + b_n . \qquad 3.13$$

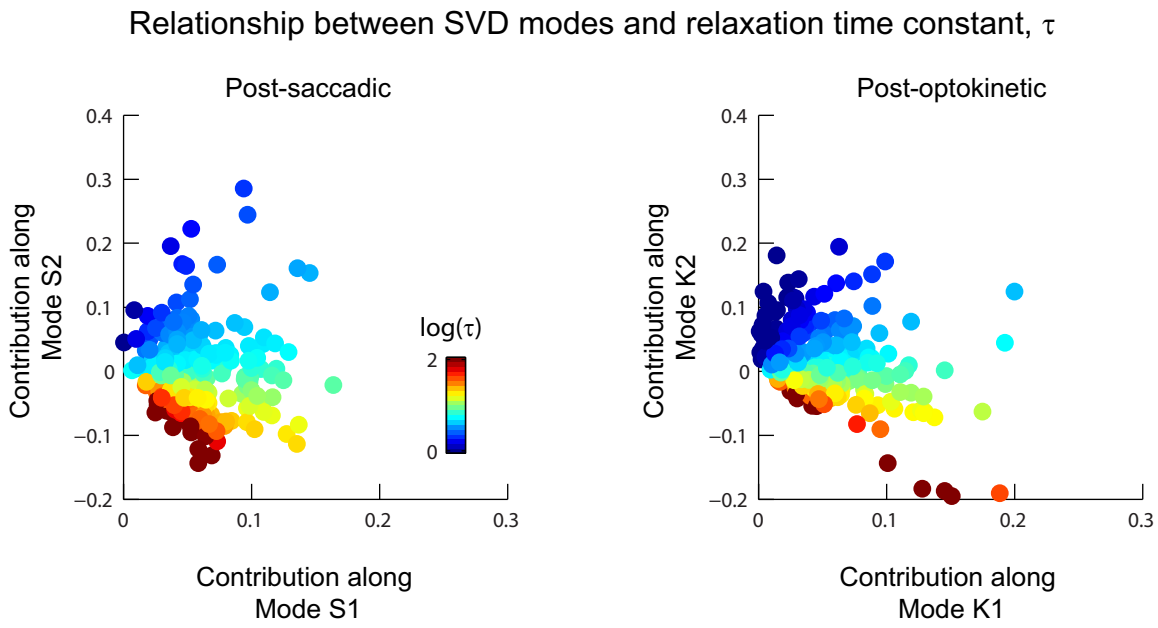
where i represents the cell number, X_i is the position of cell i along the model rostro-caudal axis, and a_n and b_n are the slope and intercept characterizing the gradient. The slopes a_n were taken from the corresponding gradients in the experimental modes $J1^*$ to $J4^*$ (Supplementary Figure 8). The intercepts were determined by solving Equation 3.13 for b_n when using the mode entries and positions associated with cells in all fish. For the model, the rostro-caudal positions X_j were then determined by using a linear regression to simultaneously solve Equation 3.13 for all four modes.

3.6 Supplementary Material



Supplementary Figure 1: SVD of firing rates during fixations.

We used the signal determined from stimulus and fixation related fluorescence data to identify the modes related to persistent firing during the fixation period. As described in the Methods, we first determined the three components of the significant SVD fluorescence modes identified in Figure 3.3, and then deconvolved to obtain the firing rate signal spanning both the stimulus and fixation periods. A second SVD was then performed on the resulting firing rates during the fixation period in order to determine the minimum number of modes needed to account for the fixation-related firing; the firing rate projections along the resulting modes are shown in **a**. In this plot we see that, during the fixation period, the persistent firing rates along the third mode are very small in both cases. To quantify this observation, we calculated the fraction of the total variation relative to both zero (γ^{zero}) and the mean (γ^{mean}) for each component (equations 3.8 and 3.9, respectively) as shown in **b** and **c**. We found that greater than 95% of the mean and variance in firing is contained along the first two modes during both post-saccadic and post-optokinetic fixations.



Supplementary Figure 2: Relationships of SVD modes to persistence time constant.

The SVD analysis of the main text revealed that persistent activity during either post-saccadic or post-optokinetic fixations could be understood as arising from a linear combination of the firing rate contributions along two modes, one associated with a relatively large mean value over time and slow dynamics (Figure 3.3c, dark red), and another with a much smaller mean value and fast dynamics (Figure 3.3c, yellow). Here we examine how the relative contributions along each mode relate to the firing rate dynamics of any given neuron.

As described in the Methods (equation 3.10), the persistent firing rate r_i^{fix} of a given cell i is determined by a weighted linear combination given by

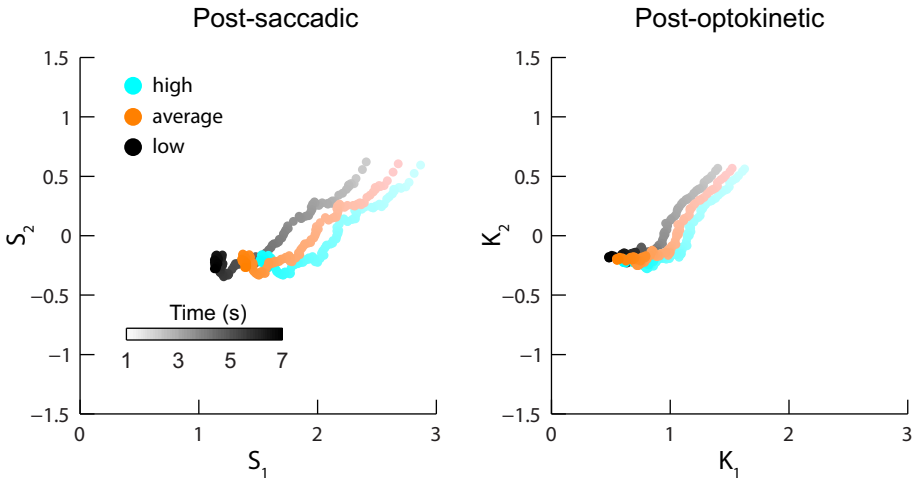
$$r_{t,i}^{fix} = V_{i,1}^{r} C_{1}^{r}(t) + V_{i,2}^{r} C_{2}^{r}(t),$$

where the components $C^r(t)$ indicate the time-varying contributions to population-level activity along each mode, and the mode entries V_i^r indicates the weight of the contribution to the firing rate of cell *i*. To examine how the components along each mode contributed to the time constant of a given cell, we generated a scatter plot of the component weights for all recorded cells, and labeled the data point for each cell with a color indicating the logarithm (base 10) of that cell's persistence time constant. Persistence time constants τ were calculated by fitting the firing rates given by the equation above with an exponential of the form $ae^{-t/\tau}$.

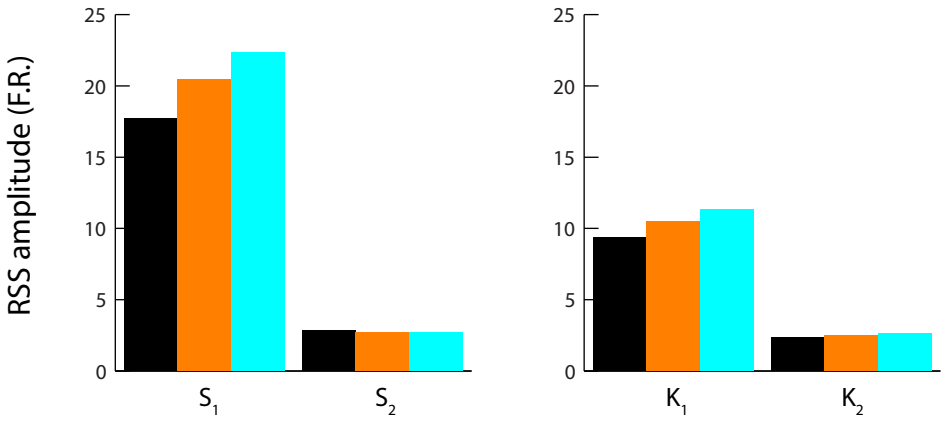
For both post-saccadic (left) and post-optokinetic (right) fixations, we observed a radial pattern of spokes along which the time constant was unchanged, indicating that the relative weighting of the first and second components of activity was the key determinant of relaxation time. For a given contribution from component 1, the contribution from component 2 was almost absent for those cells with time constants near the median (green points); relatively large and positive for those cells with the shortest time constants (blue points); and relatively large and negative for cells with the longest time constants (red points). Thus, the first component can be thought of as providing the component of activity corresponding to the dominant time scale of

relaxation across the population, and the second component as providing the changes around this median time scale that establish the heterogeneity in relaxation times.

a Firing rate trajectories during fixations at **high** and **low** eye positions

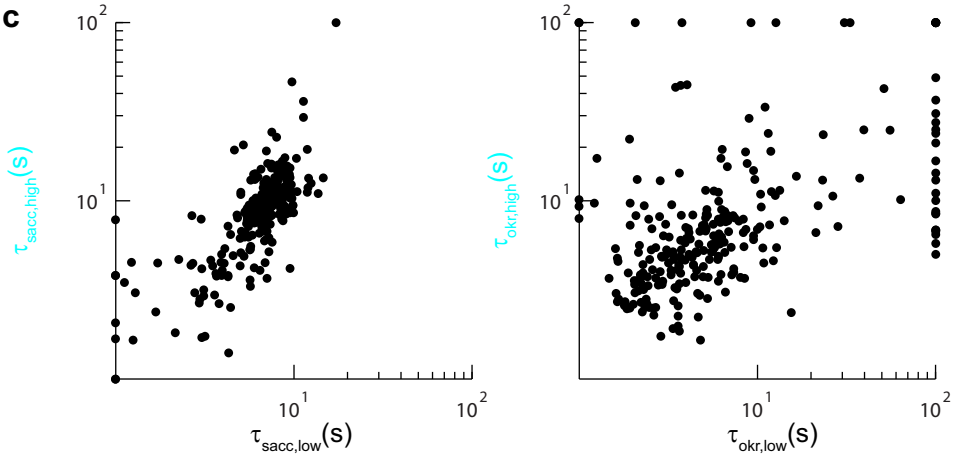


b Firing rate amplitudes during fixations at **high** and **low** eye positions



Persistence time constants at **high** and **low** eye positions

c

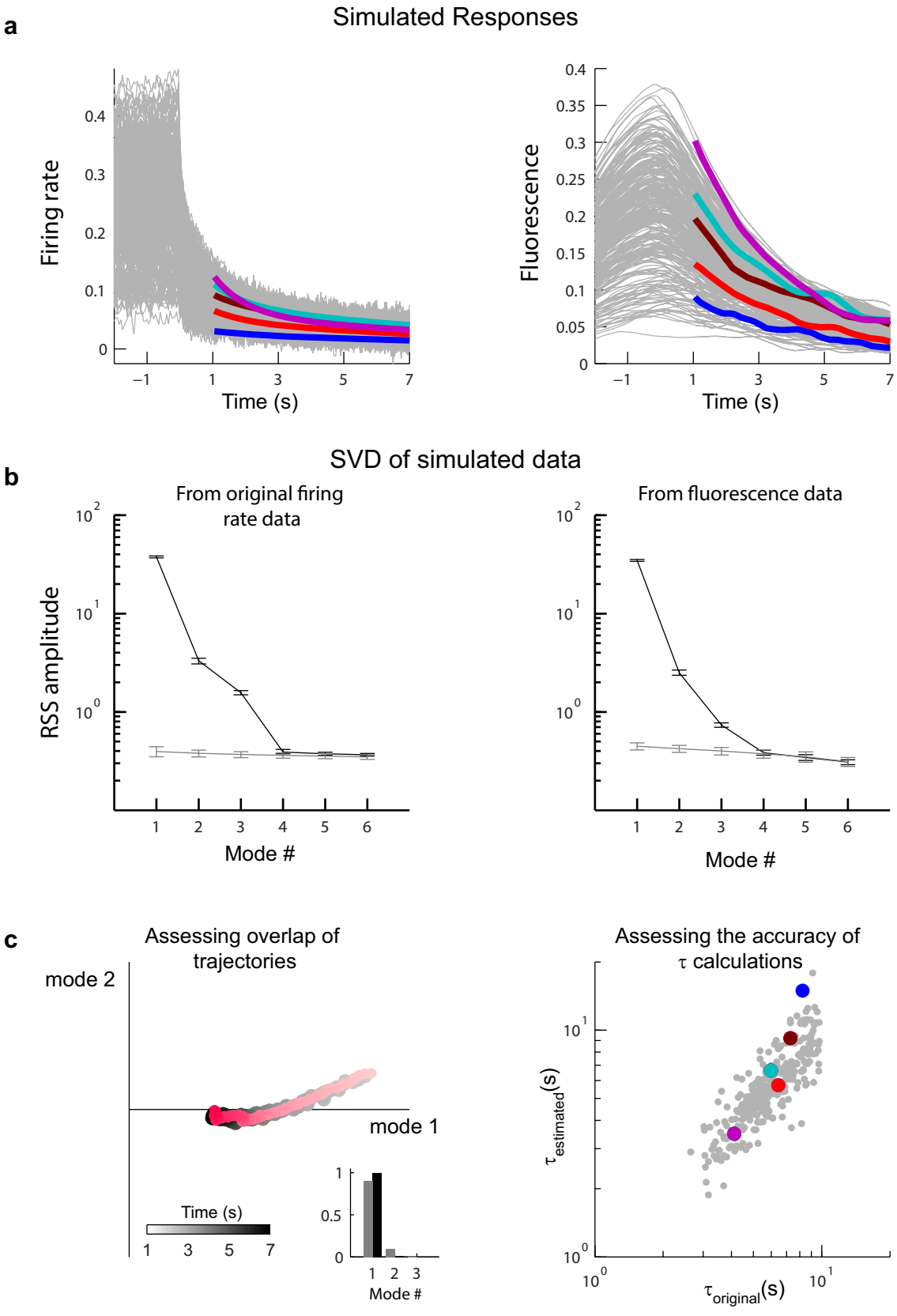


Supplementary Figure 3: Firing trajectories at different eye positions.

We examined if the patterns and dynamics of persistent firing in the VPNI varied with eye position. As presented in the main results, the firing rate of the network in each context was described by a projection along two modes, one with slow relaxation dynamics (mode 1), and the other with relatively fast relaxation dynamics (mode 2). The relative contributions of these two modes can be easily visualized by plotting the trajectory of network activity in the space identified by these two modes (**a**, orange). Here, the fast relaxation along the second mode is reflected largely in the downward trajectory at the early times, and the slow relaxation along the first mode is reflected in the leftward, more compact trajectory at later times. To explore the relationship between firing rate dynamics and eye position, for each response type we compiled two separate averages from the halves of fixations with the highest and lowest eye positions. These responses were then smoothed using the overall population's leading fluorescence components and their trajectories were plotted in the same signal space (**a**, high: cyan, low: black). These revealed trajectories of a form very similar to that of the average trajectory, with the most notable difference being translation along the primary, slowly relaxing mode. These findings qualitatively suggest that the network's dynamics, as characterized by the shape of the firing rate trajectories, do not depend on the amplitude of the velocity inputs, and that changes in eye position are largely encoded by changes along the primary mode.

To quantify these findings, we first assessed how the strength of the projections onto each mode varied with eye position. We found little change in the rss amplitude of the component along mode 2, and notable increase in the rss amplitude of the component along mode 1 with increasing eye position (**b**, lowest, average, and highest eye positions for the post-saccadic context were 6.7°, 7.3° and 8.3°, and for the post-optokinetic context were 6.7°, 7.1°, and 8.5°). Together, these findings quantitatively support the idea that changes in eye position are represented primarily by changes along the primary mode.

Next we compared, for each cell, the relaxation time constant during fixations at high and low eye positions (**c)**. For post-saccadic (Spearman rank correlation (src)=0.71, $p=1.9\text{x}10^{-41}$) and post-optokinetic (src=0.52, $p=7.1\text{x}10^{-20}$) fixations, the time constant distributions were highly correlated, comparable to the correlation between two randomly chosen halves of post-saccadic (src=0.66) and post-optokinetic (src=0.67) fixations. These results demonstrate that the firing dynamics, as characterized by their relaxation time constant, do not depend on eye position.


a
Simulated Responses
Firing rate
Time (s)
Fluorescence
Time (s)
b
SVD of simulated data
From original firing rate data
RSS amplitude
Mode #
From fluorescence data
Mode #
c
Assessing overlap of trajectories
mode 2
mode 1
Time (s)
Mode #
Assessing the accuracy of τ calculations
τ estimated (s)
τ original (s)

Supplementary Figure 4: SVD of simulated stimulus-triggered average fluorescence.

Here we present an analysis of simulated data in order to characterize the accuracy of our SVD procedure in identifying the underlying modes of persistent firing. The arguments hold for both saccadic and optokinetic inputs, and are therefore only presented for the optokinetic case. In the following, we first describe how our simulated fluorescence data set with known mode structure is constructed. We then apply the SVD analysis of the main text to this simulated data to generate estimates of the dynamical modes in the underlying firing and show that they accurately reflect the original mode structure. We also measure the persistence time constants from the original and estimated firing rates and show that they are well correlated.

Simulated average fluorescence data was generated using realistic firing rate dynamics, noise, and data acquisition sampling rates. Firing rates were described by the relationship

$$r_i(t) = \xi(t) + \int_{-4}^{7} I(t-t')\left(a_i e^{-(t'+7)/10} + b_i e^{-(t'+7)/1} + c_i e^{-(t'+7)/.1}\right) dt' \qquad 3.14$$

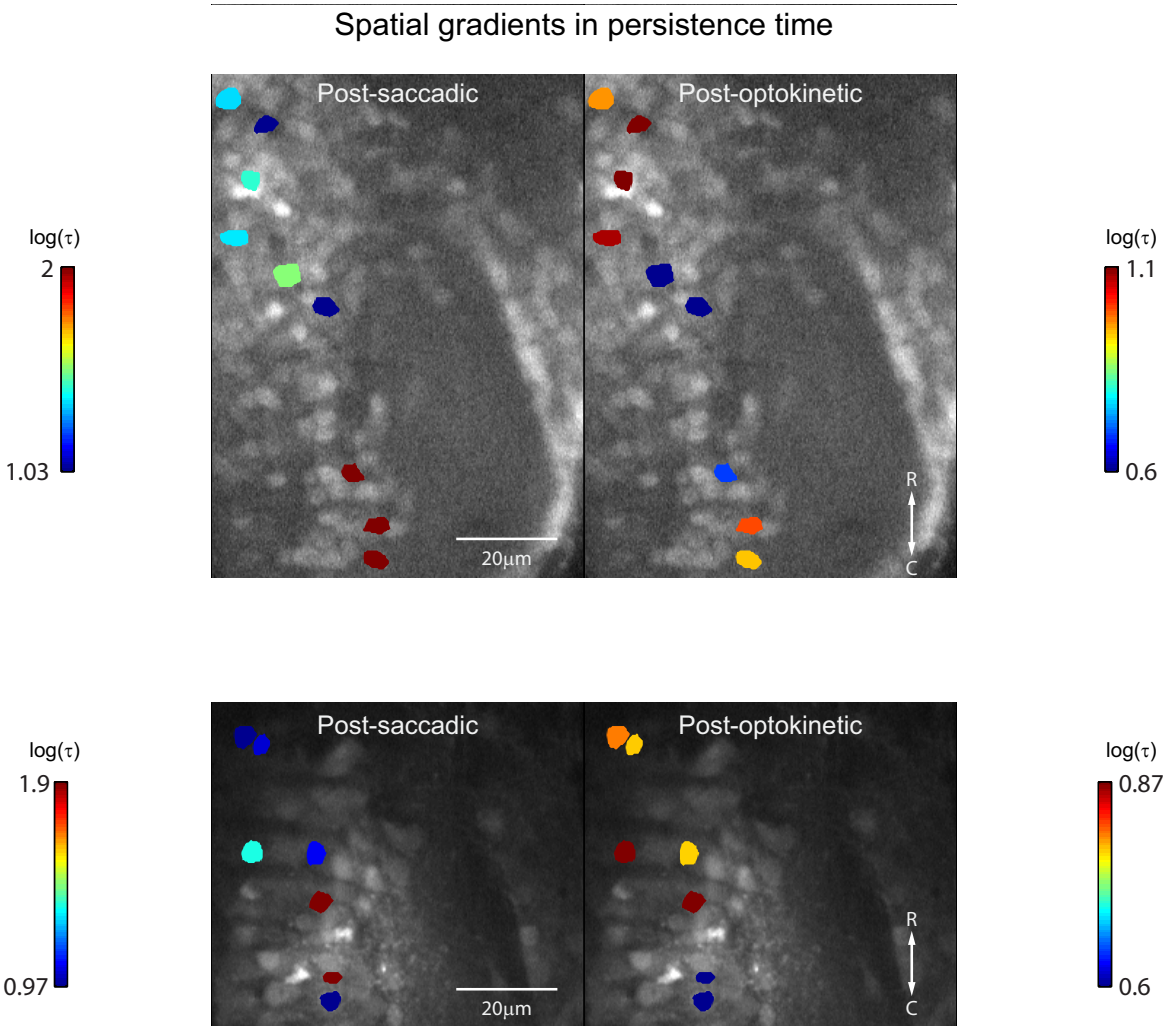
where I describes the stimulus $\xi(t)$ represents Gaussian random noise, and the exponentials describe the dynamics of firing during the fixation period. The values of a, b, and c were randomly chosen over the ranges (0.05, 0.2), (0, 0.2) and (0, 0.2), respectively. The stimulus function was equal to 1 for all times before t=0 and 0 for

all times after t=0. The noise $\xi(t)$ was specified by a mean of zero and an rms value that was approximately 25% of the average rms value of $r_i(t)$. We show in panel (**a**, left) the firing rates for all simulated cells during one fixation trial, and highlight, for the fixation window, data associated with 6 of these cells; data corresponding to these cells are also highlighted in panels (**a**, right) and (**c**, right). To generate simulated fluorescence data, we then convolved these firing rate time series with a 1.9 second filter representing the previously assessed dynamics of calcium buffering[50], added additional Gaussian-random noise with variance chosen to match the data, and then sub-sampled to 1 Hz to simulate realistic data acquisition rates. These model fluorescence values were then analyzed in a manner identical to that of the real data, interpolating to 20 Hz and then, for each cell, averaging together activity from seven trials (**a**, right), yielding simulated fluorescence data comparable to the experimentally observed post-optokinetic fixation average fluorescence.

Next, we examined how well we could determine the original firing rate mode structure when beginning our analysis with an SVD of the stimulus-triggered average fluorescence. As expected from the presence of three distinct temporal components in equation 3.14, an SVD of the original firing rate data over the -4 to 7 second window resulted in the identification of three modes along which activity was significant, as assessed by identifying those cases where the confidence intervals of the signal (black) and noise (gray) singular values did not overlap (**b**, left). The SVD of the

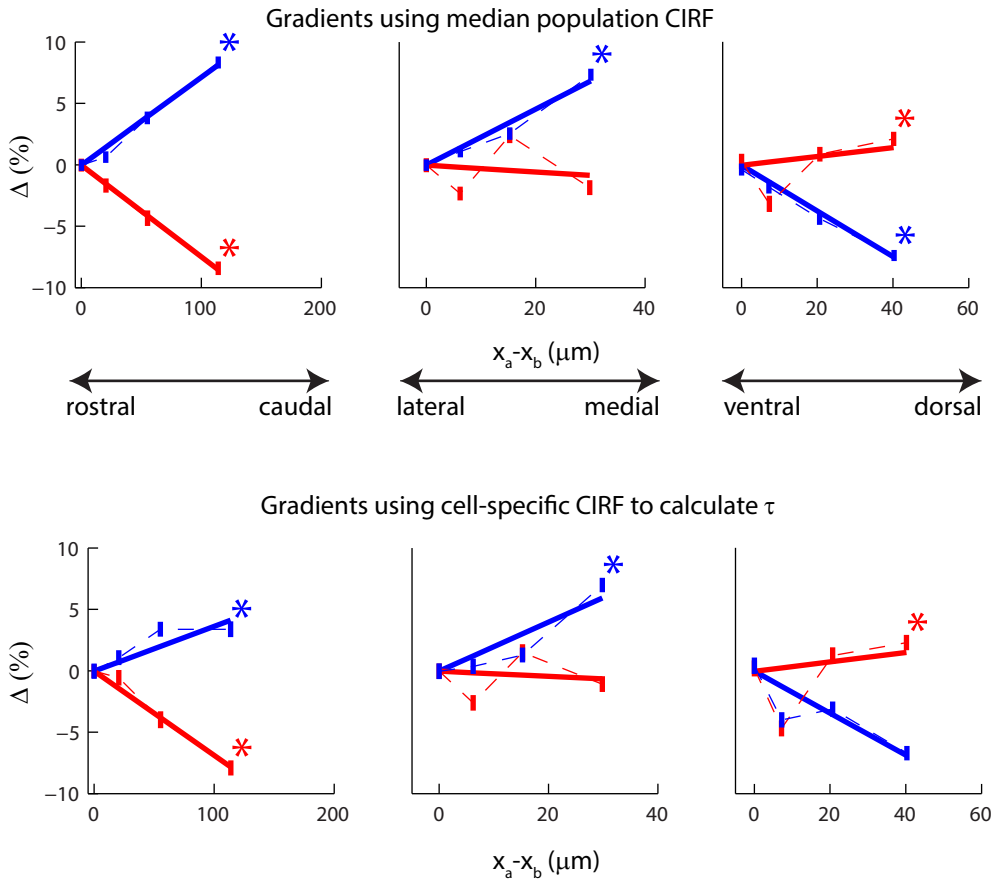
resulting fluorescence signal and noise also revealed the presence of three significant modes (**b**, right).

We next deconvolved the fluorescence signal captured by the significant fluorescence modes to determine the firing rate signal. While three firing rate modes were used when considering the entire time window, only two modes were needed to account for 95% of the mean (black) and variance (gray) of firing during the fixation period ($t \geq 1\text{s}$), as expected from the relatively fast time scale (0.1 seconds) associated with the third exponential of the firing rate dynamics specified by equation 3.14 (**c**, left, inset). The data trajectory along these two significant modes (red) matched the original trajectory (black) almost exactly. Furthermore, we found that 99.9% of the estimated first mode, and 85.3% of the estimated second mode, overlapped with the original two firing rate modes. Simulations run without noise revealed the small disagreement between these modes to be entirely due to the random noise added to the firing and fluorescence signals (data not shown). Finally, we also found that the persistence values of the individual cells assessed from single exponential fits to the original and estimated firing rates were closely correlated (**c**, right; Spearman rank correlation = 0.86).



Supplementary Figure 5: Spatial maps of persistence time constants.

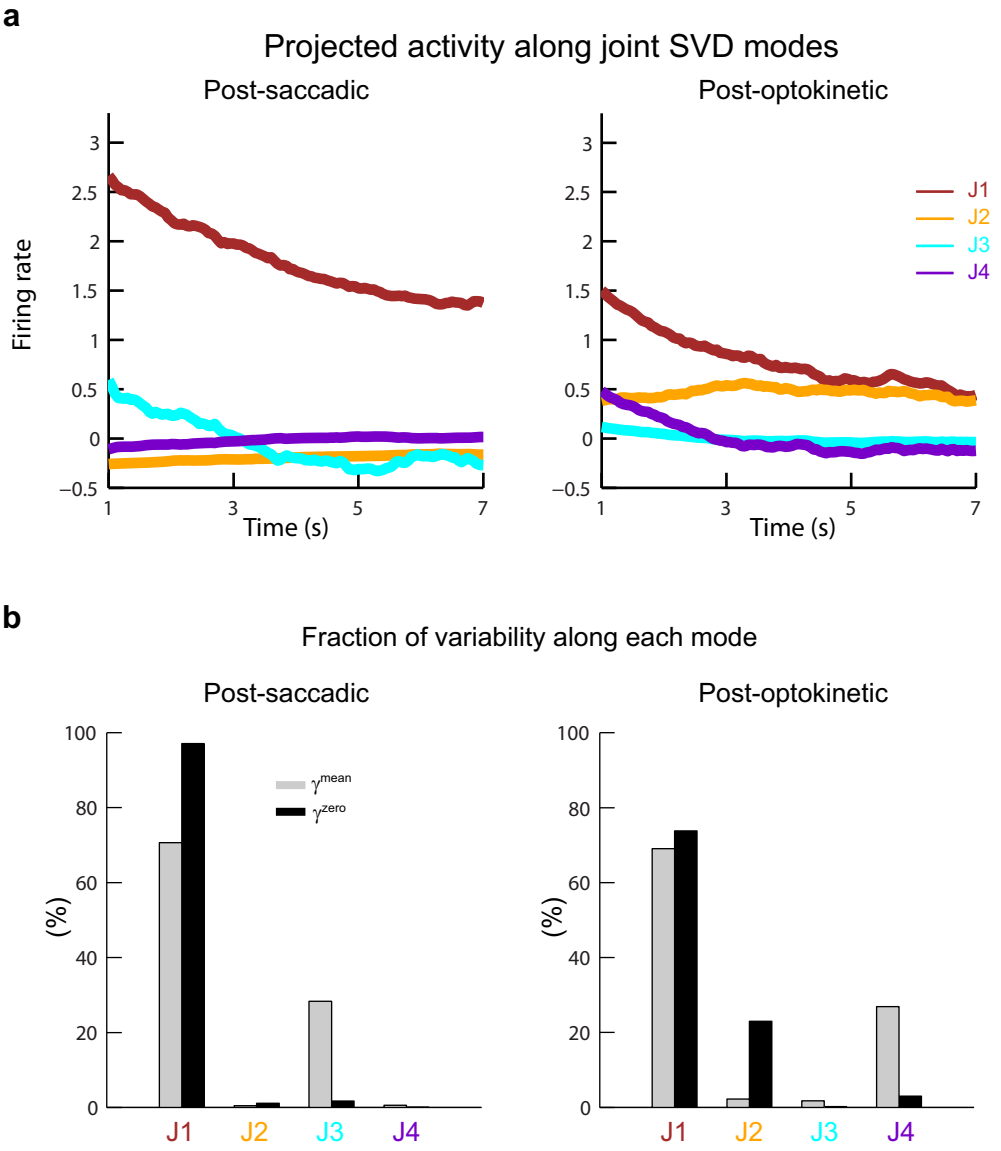
As in Figure 3.5a, here we show, for completeness, additional average images within a single optical plane where VPNI cells were color-coded according to $\log_{10}(\tau)$. Note that cells with time constants outside the indicated range were coded with the color of the closest limit. On average, both examples illustrate the finding that the most rostral cells tend to be most persistent following optokinetic stimulation while the most caudal cells tend to be most persistent following saccadic stimulation.



Supplementary Figure 6: Gradients in persistent firing.

Here we assess gradients in persistent firing along all three principle spatial axes, with firing rates calculated from deconvolutions using either average (**a**) or cell-specific (**b**) values of the calcium impulse response function (CIRF; Methods). Spatial gradients in persistent firing were assessed along the rostro-caudal (left), medio-lateral (middle), and dorso-ventral (right) axes by measuring the pairwise differences of the persistence time constant, Δ (equation 3.11), versus distance between cells. Error bars represent the standard error of the mean, and asterisks denote gradients that were significantly different from zero ($p<0.05$). For gradients calculated using the average

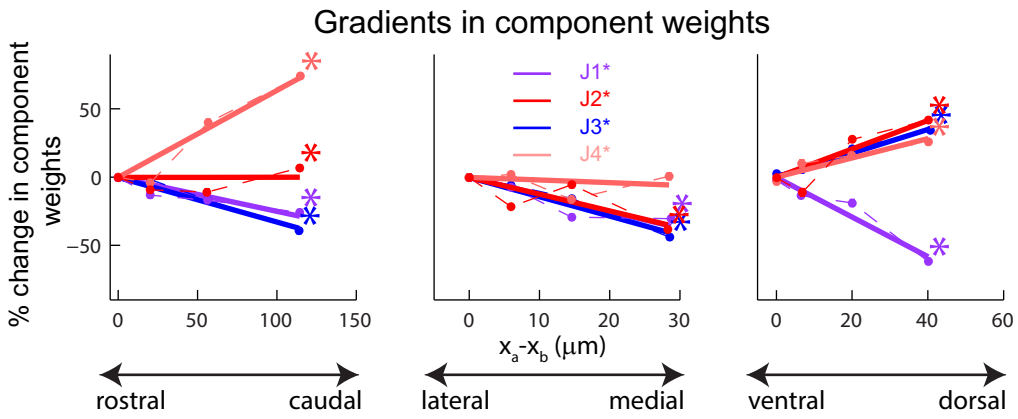
CIRF, src values along the rostro-caudal, medio-lateral, and dorso-ventral axes were, respectively: 0.13 ($p=7.1x10^{-19}$), 0.13 ($p=1.9x10^{-17}$) and -0.09 ($p=1.1x10^{-7}$) during post-saccadic fixations, and -0.12 ($p=1.4x10^{-15}$), 0.01 ($p=0.44$) and 0.07 ($p=1.3x10^{-5}$) during post-optokinetic fixations. Similar results were obtained when using cell-specific values of the CIRF time constant. Along the rostro-caudal, medio-lateral, and dorso-ventral axes, src values were, respectively: 0.04 src($p=7.0x10^{-3}$), 0.09 ($p=3.2x10^{-10}$) and -0.03 ($p=0.08$) during post-saccadic fixations and -0.11 ($p=1.3x10^{-13}$), 0.01 ($p=0.46$) and 0.10 ($p=1.2x10^{-8}$) during post-optokinetic fixations. In both cases, gradients significantly differed between contexts ($p<10^{-5}$) along all three axes, with the most notable differences always along the rostro-caudal axis. The presence of gradient differences along more than one spatial dimension is consistent with a lack of correlation in the time constants of fixation dynamics across contexts (Figure 3.3c).



Supplementary Figure 7: Joint SVD components

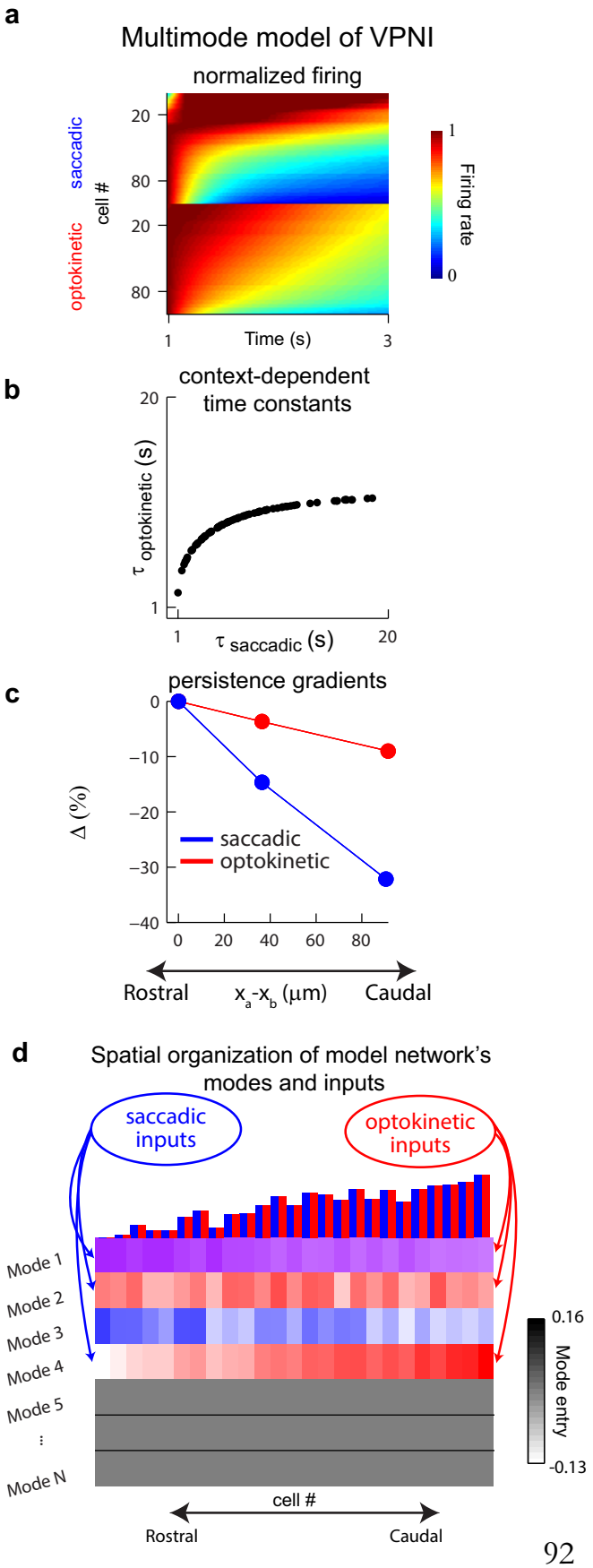
(**a**) When viewed together, the post-saccadic and post-optokinetic signals were contained within a space defined by four modes, joint modes J1 through J4. For both contexts, the components along the first mode had relatively large magnitudes and slow dynamics. The second component was activated differently for each context, with a small negative amplitude during post-saccadic fixations and larger positive

amplitude during post-optokinetic fixations. The third and fourth components had zero mean and faster dynamics, with the third component prominent during post-saccadic fixations and the fourth prominent during post-optokinetic fixations. (**b**) As in Supplementary Figure 1, we calculated the fraction of the total variation relative to both the mean (γ^{mean} ; gray) and zero (γ^{zero} ; black) for each component, J1-J4. In the plot we see that all four modes were required to explain at least 95% of both γ^{zero} and γ^{mean} .



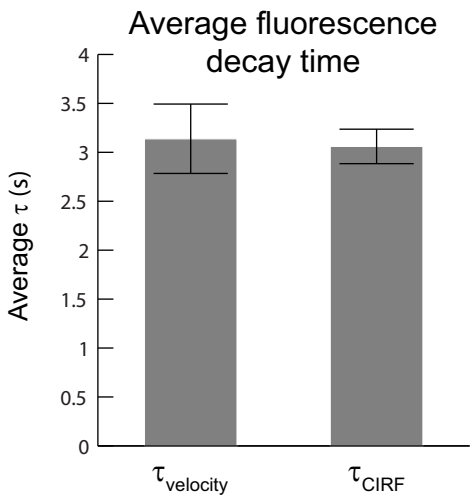
Supplementary Figure 8: Spatial gradients in Modes J1* through J4*.

Here we characterized any potential spatial organization among the modes J1*-J4*. For each mode, gradients were calculated by taking for all pairs of cells pairwise differences between the mode's entries (i.e. $V_{i,n}$ in equation 3.2), normalized by the population-averaged absolute pairwise difference, and comparing these differences to the pairwise distances between the cells along the rostro-caudal (left), medio-lateral (middle), and dorso-ventral (right) axes. We found significant ($p<0.05$) trends along all three axes (indicated by asterisks).


a
Multimode model of VPNI
normalized firing
saccadic
optokinetic
cell #
20
80
20
80
1
Time (s)
3
1
Firing rate
0
b
context-dependent
time constants
20
τ optokinetic (s)
1
1
τ saccadic (s)
20
c
persistence gradients
0
−10
−20
−30
−40
Δ (%)
saccadic
optokinetic
0
20
40
60
80
Rostral
x_a-x_b (μm)
Caudal
d
Spatial organization of model network's
modes and inputs
saccadic
inputs
optokinetic
inputs
Mode 1
Mode 2
Mode 3
Mode 4
Mode 5
Mode N
0.16
Mode entry
-0.13
cell #
Rostral
Caudal

Supplementary Figure 9: Context-independent activation of simulated network.

We sought to determine if stimulating the four-mode network specified by our data in a manner independent of context could similarly replicate the heterogeneity and spatial gradients observed in Figures 4 and 5. Context-independent stimulation was modeled by providing the saccadic input that activated modes 1, 2, and 4 with the same pattern as optokinetic stimulation (**d**). We found that although context-independent activation was able to generate the appropriate heterogeneity in persistent firing at the population level (**a**), on a cell-by-cell level persistence values in such context-independent models were highly correlated (**b**), and spatial gradients were not reversed (**c**). Hence, our results reveal that context-independent activation is unable to reproduce the dynamics observed in the VPNI.



Supplementary Figure 10: Assessing persistent activity in velocity-storage neurons.

Here we examined if there might be any contribution to post-optokinetic fixations arising from the velocity storage system. In some species, it has been shown that the activity of velocity inputs to the integrator following sustained (10's of seconds) optokinetic stimulation does not decay to zero instantaneously, resulting in eye movements that slowly drift in the direction of the previously moving visual stimulus, a behavior referred to as post-optokinetic nystagmus[89]. This behavior is believed to be mediated by the continued firing of velocity-storage neurons that maintain a signal encoding eye-velocity commands[90]. Although previous studies of eye movements have shown that larval zebrafish do not yet exhibit velocity storage behavior[86], since this signal, if present, would provide indirect input to the VPNI during the designated fixation period, it is important to consider if velocity-storage

could have contributed to the differences in VPNI dynamics observed between post-optokinetic and post-saccadic fixations.

To examine this issue further, we characterized the dynamics of putative velocity-storage neurons during post-optokinetic fixations. Velocity-storage neurons in the larval zebrafish, also located in rhombomeres 7 and 8[91], were identified as neurons whose activity was correlated with eye movements during OKR, but not during saccades[90] (61 cells). During post-optokinetic fixations, the fluorescence of these cells decayed quickly to baseline, with a mean time constant of 3.14 ± 0.12 (sem) seconds (left bar in plot). This value did not differ significantly from the mean time constant of the calcium impulse response function (CIRF) determined from fluorescence decay times of VPNI neurons following contraversive saccades (right bar in plot 3.06 ± 0.53 seconds; p=0.11, Wilcoxon rank-sum test). This suggests that the firing rates of velocity-storage neurons rapidly drop to a baseline level during post-optokinetic fixations, consistent with previous behavioral findings[86]. Hence, the dynamics of persistent firing in VPNI neurons during post-optokinetic fixations should depend only on the mechanisms of integration within the VPNI.

Chapter 4 Inferring connectivity from population activity

4.1 Abstract

In this chapter, linear-regression is used to fit a network model to the dynamics of oculomotor integrator neurons recorded during behavior. Because of the low-dimensional nature of these data, a wide range of solutions are able to fit the data equally well. To quantify this variability, we use a genetic algorithm to randomly sample the space of best-fit solutions. Within the distribution of solutions we find that post-saccadic responses are consistent with both a symmetric and a feedforward chain of connections. However, feedforward connectivity is inconsistent with fits to both post-saccadic and post-optokinetic responses. Finally, we find that solutions with two vastly different spatial organizations exist, one in which connections are strongest between nearest neighbors, and another in which connections are strongest between more distally-located cells.

4.2 Introduction

Previous experiments in the oculomotor integrator[25–27] have implicated network interactions as the dominant mechanism underlying the maintenance of persistent firing. Recently the results of this experimental work have been incorporated into a regression-based modeling framework in an attempt to identify connectivity structures consistent with experimental observations[8]. In this chapter, we adapt this modeling framework to the study of connectivity underlying the dynamics recorded in the zebrafish oculomotor integrator (Chapter 3).

As described in Chapter 2, using low-dimensional neural activity to solve for connectivity can result in solutions which are not unique. To quantify the degree to which solutions can vary, we perform a random sampling of the distribution of best-fit solutions. We find that because of the context-dependent activity described in Chapter 3 the inclusion of both post-saccadic and post-optokinetic fixation dynamics in the fits results in a decrease in the variability in connection strengths. Finally, we search through the distribution of best-fit solutions for general features common among all solutions.

4.3 Methods

4.3.1 Linear regression on network dynamics

We approximate the dynamics of the integrator network with a linear network equation of the form[15]

$$\tau \frac{dr_{t,i}}{dt} = -r_{t,i} + \sum_{j=1}^{n} r_{t,j} w_{j,i} + T_i + h_i \qquad 4.1$$

where $r_{t,i}$ is the firing rate of the i^{th} cell at time t, $w_{j,i}$ is the strength of the synaptic connection from cell j onto cell i, T_i and h_i are the tonic and velocity inputs, respectively. Equation 4.1 can be rewritten such that

$$\tau \frac{dr_{t,i}}{dt} + r_{t,i} = \sum_{j=1}^{n} r_{t,j} w_{j,i} + T_i \qquad 4.2$$

where the left hand side represents known quantities, and the term h_i was set to zero because velocity inputs during fixations are silent. The connections $w_{j,i}$ and T_i can then be estimated using linear regression. Equation 4.2 can be rewritten in matrix form as

$$\begin{pmatrix} \frac{dr_{t=1,i}}{dt} \\ \vdots \\ \frac{dr_{t=T,i}}{dt} \end{pmatrix} + \begin{pmatrix} r_{t=1,i} \\ \vdots \\ r_{t=T,i} \end{pmatrix} = \begin{pmatrix} r_{1,1} & \cdots & r_{1,N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ r_{T,1} & \cdots & r_{T,N} & 1 \end{pmatrix} \begin{pmatrix} w_{1,1} & \cdots & w_{1,N} & T_i \\ \vdots & \ddots & \vdots & \vdots \\ w_{T,1} & \cdots & w_{N,N} & T_i \end{pmatrix} \qquad 4.3$$

or more succinctly as

$$\mathbf{Y} = \mathbf{rW} \qquad 4.4$$

This equation can be solved by inverting the matrix of firing rates **r** so that

$$\mathbf{W} = (\mathbf{r})^{-1}\mathbf{Y} \qquad 4.5$$

In order for a matrix to be invertible it must have full rank such that its columns form a linearly independent basis with *N* dimensions, where *N* is the number of columns or, in this case, cells. As seen in the previous chapter, our matrix $\mathbf{r}$, which consists of firing during both post-saccadic and post-optokinetic fixations, is comprised of only four linearly independent components. Hence, solving equation 4.4 requires the implementation of a method that produces an estimate of the inverse of a rank-deficient matrix. One such method is the Moore-Penrose pseudo-inverse. To better understand the logic behind this method we rewrite the firing rates in equation 4.4 using the SVD as $\mathbf{r} = \mathbf{usv}'$. Equation 4.4 becomes:

$$\mathbf{Y} = \tilde{\mathbf{r}}\tilde{\mathbf{w}} \qquad 4.6$$

Where $\tilde{\mathbf{r}} = \mathbf{us}$ is simply the projection of the firing rates along the SVD modes (Figure 3.3c) and $\tilde{\mathbf{w}} = \mathbf{v}'\mathbf{w}$ is a matrix of connections from these components to each cell. For example element $\tilde{w}_{i,j}$ is the strength of the contribution from the i^{th} component to the firing of the j^{th} cell. Because we know that the firing along modes 5 through N is just noise, or zero after SVD smoothing is performed, we can set the connections from these modes equal to zero. In other words, we can set rows 5 through N of $\tilde{w}$ equal to zero.

Characterization of a solution's uniqueness can be accomplished by studying the geometric properties of the cost function. For a linear least-squares fit χ_j^2, as a function of $\mathbf{W}$, takes the form of an *N*-dimensional parabola (Figure 4.2).

$$\chi_j^2 = \sum_{t=1}^{T}\left(y_{t,j} - \sum_{k=1}^{n} r_{t,k} w_{k,j}\right)^2 \qquad 4.7$$

In Figure 4.2 we see that there are some directions along which this function changes rapidly, and others along which it changes minimally. Because, for a best-fit solution, the first derivative (gradient) of χ_j^2 should be zero, χ_j^2's steepness is best-characterized by the second derivative. A matrix of partial second derivatives, or Hessian matrix, can be computed as follows

$$\begin{aligned} \frac{\partial \chi^2}{\partial w_{p,j}} &= 2\sum_{t=1}^{T}\left(y_{t,j} - \sum_{k=1}^{n} r_{t,k} w_{k,j}\right) r_{t,p} \\ \frac{\partial \chi^2}{\partial w_{q,j}\partial w_{p,j}} &= -2\sum_{t=1}^{T} r_{t,q} x_{t,p} \\ &= -2\mathbf{r}'\mathbf{r} \end{aligned} \qquad 4.8$$

We can then diagonalize this matrix to determine the directions along which the cost function is most sensitive to changes in *w*. A given eigenvalue reveals the steepness of the cost function along the direction specified by its corresponding eigenvector. Diagonalization of $\mathbf{r}'\mathbf{r}$ yields

$$\mathbf{r}'\mathbf{r} = \mathbf{V}\mathbf{D}\mathbf{V}' \qquad 4.9$$

Interestingly, if we rewrite 4.9 in terms of the SVD we get

$$\begin{aligned} \mathbf{r'r} &= (\mathbf{USV'})'(\mathbf{USV'}) \\ &= \mathbf{VSU'USV'} \\ &= \mathbf{VS^2V'} \end{aligned} \qquad 4.10$$

And therefore, the eigenvalues of the Hessian matrix are equal to the square root of the singular values of the firing rate matrix $\mathbf{r}$. Since the firing rate is only non-zero along four modes this means that the fits are insensitive to movements along the other N-4 modes in the system.

4.3.2 Exploring space of best-fit solutions

Because of the low-dimensional nature of the fixation-related firing described in Chapter 3, equation 4.8 implies that there must exist N-4 directions in solution space along which the fits are insensitive. In a situation in which the connection strengths are not constrained, alterative solutions can be obtained simply by moving along these insensitive directions. However, previous results have suggested that the connections between ipsilateral pairs of neurons in the oculomotor integrator must be positive[26,27]. By placing these hard boundaries in the space of solutions, it is possible that directions that were previously deemed insensitive can no longer be traversed due to the presence of the boundaries, therefore, a modified approach is required to identify the space containing the best-fit solutions. Furthermore, we restrict the maximal connection strength such that $W_{max} = \frac{1}{5}\frac{50}{N}$, were N is the number of cells recorded from in a given fish. This value was chosen to ensure that, on average, a

given post-synaptic neuron in a network of 50 cells would require input from at least 5 other cells. Here we employ a genetic algorithm (GA) which is a stochastic method for solving optimization problems to generate a population of randomly distributed solutions of equal fit quality.

To begin a starting solution, W_0 composed of all positive weights is determined by using a Nelder-Mead simplex algorithm (Matlab; lsqCurveFit). The fitness of each solution was then assessed by computing, for the j^{th} cell,

$$R_{j,0}^2 = 1 - \frac{\sum_i (y_{i,j} - r_{i,k} W_{k,j})^2}{\sum_i y_{i,j}^2} \qquad 4.11$$

This value is then used as a reference for all candidate solutions; if R^2 for a particular solution is within 99.9% of $R_{j,0}^2$, that solution is deemed part of the population of best-fit solutions. By ensuring that the fit quality of all solutions are sufficiently close to that given by equation 4.11 we ensure that all differences in connection weights occur along directions in which the signal amplitude is insignificant. Ideally then we should require that any candidate solution's R^2 to be equal to $R_{j,0}^2$. However, using the approach outlined below using a threshold value of 100% has proven computationally prohibitive. A limited subset of the data in which solutions were required to have a fit quality equal to $R_{j,0}^2$ had results that had the same qualitative features as those obtained by requiring fit quality to be within 99.9% of $R_{j,0}^2$.

As mentioned above, obtaining solutions with fit qualities that were within 99.9% of $R^2_{j,0}$ were obtained with a genetic algorithm. A genetic algorithm is a procedure in which a population of randomly chosen solutions iteratively "evolves" towards solutions with a pre-specified level of fitness. Because the initial population of solutions are randomly chosen and because the population is subject to random mutations during the evolution, the final set of best-fit parameters are randomly distributed throughout parameter space. The initial population used here consisted of 150 randomly chosen perturbations of the best fit solution, W_0, obtained through solving equation 4.4 subject to the constraint that all parameters be positive. For the initial population there perturbations could be as large as 5 times larger than W_0. Fits for each cell were performed individually. The fitness of each solution was then assessed by computing, for the j^{th} cell, $R^2_{j,\text{gen}}$. Pairs of solutions were then randomly chosen from the population, and the solution with the highest fitness was put into a "breeding pool". After competing in this "tournament selection" the pair of solutions was also put back into the overall population and the procedure was repeated until the breeding pool contained 150 solutions. Pairs of solutions were then selected from this pool and "bred" together. The i^{th} parameter of the resulting offspring was equal to the i^{th} parameter of one of the parents chosen at random with equal probability. Each breeding pair produced two offspring. Offspring were then subjected to random mutations with a 1% chance that that for each parameter to change by a randomly

selected magnitude. This process was then repeated with this new population taking the place of the previous. The process was iterated until 1 of three criteria was reached: 1.) 20 members of the population had R^2 values that were at least as large as $0.999\,R^2_{j,0}$; 2.) 100 generations were produced without reaching this threshold; 3.) The largest and smallest R^2 values in the population differed by less than 0.02; in other words, the population evolved to a steady state without reaching threshold fitness. If the first criterion was met, the 20 solutions with the best fitness were stored and the process was repeated N (number of cells) times in total yielding N x 20 solutions. If criteria 2 or 3 were met first, the evolutionary process was repeated. Each time the process had to be repeated after failing to reach the threshold level of fitness the system was initialized closer to the starting solution W_0 until after 20 such iterations the GA solution was set equal to W_0. In such an instance, because the search failed to find solutions which were of within 99.9% of $R^2_{j,0}$, it was determined that W_0 was essentially a unique solution for that particular cell.

4.3.3 Quantifying spread in solutions

Once the population of solutions was obtained for the genetic algorithm we computed the spread in parameter values. This was done first by sorting all possible values of each connection $W_{i,j}$ and computing, W_{range} the difference between the values whose magnitudes were in the 97.5 and 2.5 percentiles. To compare the results

obtained from different animals with differing numbers of recorded cells, we normalized W_{range} by the maximal connection strength $W_{\max}$. Next, we quantified the number of insensitive directions that remained after adding the positivity constraint. To do this, we formed an N x (N x 20) matrix of all GA connections onto a given cell where each column represents an individual GA solution. We then subtracted the initial values $W_{0,j}$ from each column and computed an SVD of the resulting connection weights. The i^{th} singular value specified how far the solution could differ from $W_{0,j}$ along the i^{th} direction without changing the fit quality. As with W_{range}, these singular values were normalized by the maximal weight $W_{\max}$.

4.3.4 Searching for structure in GA solutions

The overall goal of the GA analysis was to determine if the structural features of the connectivity matrix suggested by the results of Figure 3.5 were in fact present in the network. More specifically, we wanted to determine if the connections were strongest between neurons with similar time constants and spatial locations. To do this we sorted each GA solution based on the slope of the relationship between the connection strength between two cells and the difference in their respective time constant rank and also based on the slope of the connection strength vs. pairwise distance relationship. This type of sorting enabled us to quantify the degree to which these two relationships could vary without changing the quality of the fits.

4.4 Results

To begin, we show solutions to equation 4.4 obtained by fitting only post-saccadic dynamics using the Moore-Penrose pseudoinverse. We then characterize the structure of the resulting solution and find that connections are strongest between cells with similar time constants and similar spatial locations. Next, we sample the space of best-fit solutions to assess the degree of variability in the solutions, and characterize the degree to which this variability decreases when post-optokinetic dynamics are included in the fits. Finally, we explore the distribution of best-fit solutions to find that, in all solutions, connections are strongest between cells with similar time constants, but the strongest connections need not be between neighboring neurons.

4.4.1 Linear network fits during post-saccadic fixations

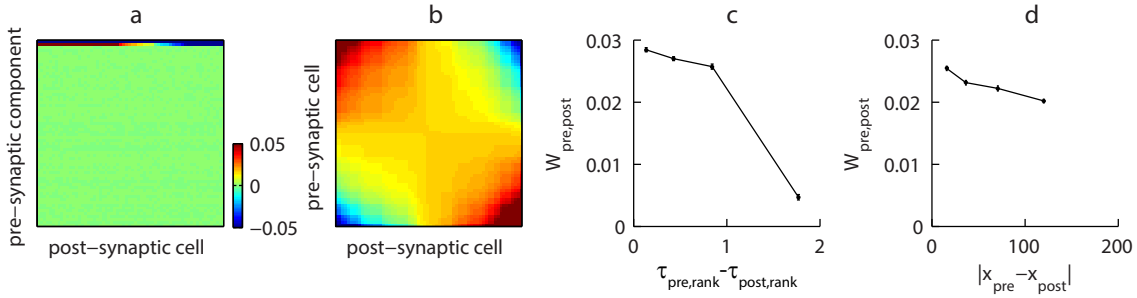


Figure 4.1: Example connectivity matrix**.** a.) Connections from modes onto post-synaptic cell ($\tilde{W}$). b.) Connections between cells. c.) Relationship between connection strength and difference in τ_{rank}. d.) Relationship between connection strength and spatial distance between cells.

The solution to equation 4.4 using the Moore-Penrose pseudo-inverse while fitting post-saccadic dynamics is shown in Figure 4.1. As described in the Methods, the pseudo-inverse in the firing rate matrix $\mathbf{r}$ can be accomplished by setting the contributions from modes along which activity is not significant equal to zero. Figure 4.1a shows the matrix of connection strengths, $\tilde{\mathbf{W}}$, from all modes (rows) onto all cells (columns) obtained by fitting equation 4.6 with firing rates measured during post-saccadic fixations from a single fish. Here we see that the only non-zero connections are those arising from the first two modes in the network. Taking the product $\mathbf{V}\widetilde{\mathbf{W}}$ yields the connectivity matrix $\mathbf{W}$ shown in Figure 4.1. Element $w_{i,j}$ in this matrix represents the connection from cell *i* onto cell *j* with the cells arranged according to average time constant (τ_{sacc}) with the top row representing connections from, and left column representing connections onto, the least persistent cell. The concentration of strong connections near the diagonal of this matrix suggests that cells with similar time constants tend to connect more strongly. In Figure 4.1-c, we see that, for this particular connectivity matrix, the connection strength between two neurons falls off as a function of the difference between their respective time constants (src=-0.42 and p=0). Because we have observed pairwise difference in time constant to vary as a function of distance between neurons, we next checked whether there was a relationship between connection strength and distance between neurons.

As illustrated in Figure 4.1-d, we observed an inverse correlation between inter-cellular distance along the rostro-caudal axis and connection strength (src=-0.09 and $p=2.3\text{x}10^{-17}$). These results provide support for the idea that connections are structured according to both difference in time constant, as well as spatial separation, intuitively suggested by the spatial gradients in time constants shown in Figure 3.5.

4.4.2 Assessing the uniqueness of best-fit solutions

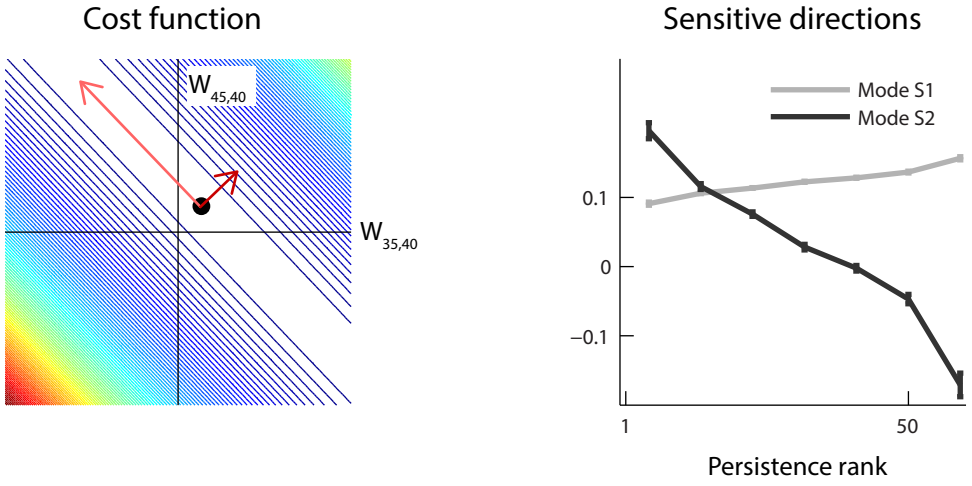


Figure 4.2: Cost-function. a.) Cost function for connections from cells (ordered by τ) 35 and 45 onto cell 40. In this plane, the sensitive and insensitive directions are illustrated by the dark and light red arrows, respectively. b.) Mode entries for sensitive directions as a function of τ_{rank}.

Because of the rank-deficient nature of our firing rate matrix $\mathbf{r}$ we know that the solution shown in Figure 4.1 is not unique. This degeneracy in the best-fit solutions to equation 4.4 can be understood by considering the matrix $\widetilde{\mathbf{W}}$. Because the activity along modes 3 through N in the network were determined to be

dominated by noise (Chapter 3), the firing rates along these modes were set to zero during the smoothing step. Hence, the connection strengths from these modes represented by rows 3 through N can be modified arbitrarily without altering the quality of the fits. The insensitivity of the fit quality is illustrated in Figure 4.2a which shows a 2D contour plot of the N-dimensional cost function given by equation 4.7 vs. the connection from two arbitrarily selected cells. The light and dark red arrows in this figure illustrate the sensitive and insensitive directions, respectively. We see that movement along the insensitive direction corresponds an increase in the connection from one cell and a corresponding decrease from the other, while movement along the sensitive direction corresponds to increasing the weights from both cells simultaneously. Figure 4.2 shows the mode entries for the two sensitive directions as a function of $\tau_{sacc.}$. In each case a significant trend is observed. The fact that all entries in the vector corresponding to the first sensitive direction are positive means that movement along this direction corresponds a decrease in fit quality resulting from an increase in connections from all cells simultaneously. The second sensitive direction has positive entries for cells with short values of τ_{sacc} and negative entries for cells with long τ_{sacc}. Therefore, movement along this direction corresponds to increasing connections from cells with short τ_{sacc} while simultaneously decreasing connections

from cells with long τ_{sacc}, or vice-versa. Such movement would have the effect of altering $\tau_{sacc.}$, hence lowering the quality of the fit.

4.4.3 Combining Post-optokinetic responses with post-saccadic responses

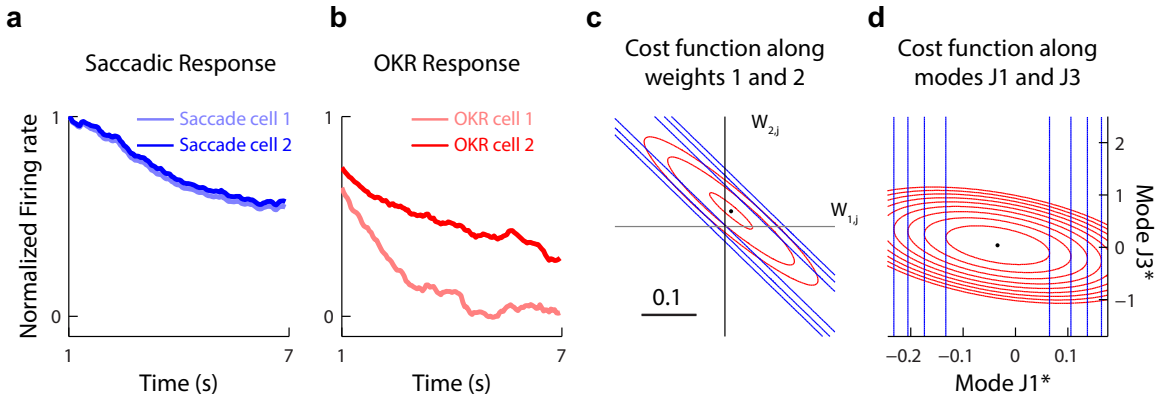


Figure 4.3: Impact of adding optokinetic data to cost function. a.) Plot of two cells whose firing rates during post-saccadic fixations were highly correlated. b.) Firing rates of same two cells during post-optokinetic fixations. c.) Cost function for fits to post-saccadic (blue) and combined post-saccadic and post-optokinetic (red) responses vs. connections from the two cells shown in a and b. d.) Cost function along modes J1* and J3*.

In the previous section we saw that for a linear fit, the sensitive directions in the cost function are the same as the dimensions along which the signal is large. Because a distinct set of modes becomes active in the integrator network during optokinetic eye movements, addition of the responses during post-optokinetic fixations should serve to better constrain the fits. This point is illustrated in the cost function in Figure 4.3. In this figure we show fit quality as a function of two cells whose activity is highly correlated during post-saccadic fixations (Figure 4.3-a), but much less so during post-optokinetic fixations (Figure 4.3-b). Because of this lack of

correlation between the post-optokinetic responses, we see that the direction along which fits to saccadic data alone were insensitive (Figure 4.3-c; blue) are now much more sensitive (Figure 4.3-c; red). When viewed as a function of modes J1* and J3* (**section**), we see that combined post-saccadic and post-optokinetic fits are sensitive along both directions while fits to post-saccadic responses along were insensitive along J3* due to the lack of post-saccadic signal along this direction.

4.4.4 Exploring parameter space using a genetic algorithm

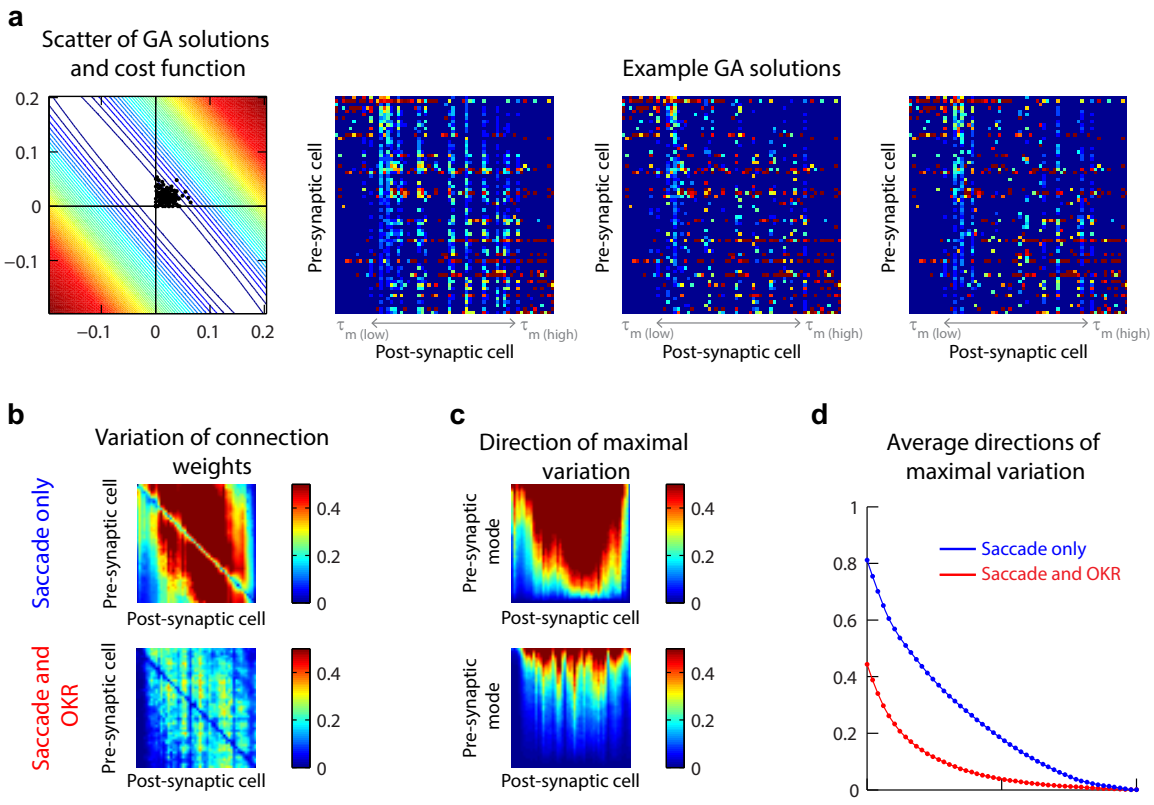


Figure 4.4: Analyzing variation in connection strengths. a.) Cost function with randomly distributed GA solutions illustrated in black (left). Three example solutions are shown in the three panels to the right. Both pre- and post-synaptic cells are arrange according to $\langle\tau\rangle$ b.) Variation in connection weights relative to the W_{max} for fits to post-saccadic responses (top) and combined post-saccadic

and post-optokinetic responses (bottom). c.) Illustration of the directions along which solutions exhibit maximal variation. d.) Average of all columns in c.

Anatomical and physiological studies[24,26,27] have suggested that connections between integrator neurons within the same side of the brain should be strictly excitatory. In other words, the entries in the connectivity matrices fit to these data should be all positive. Constraining fits to be positive complicates the exploration of solution space. This complication is illustrated in Figure 4.4-a. This figure illustrates the presence of an insensitive direction which intersects with the x and y axes. Because weights are constrained to be positive, movement along the insensitive direction in constrained to be within the first quadrant.

Random sampling of solutions with all-positive connection weights along the insensitive directions was accomplished using a genetic algorithm (GA; Methods). Solutions obtained with this approach are plotted in Figure 4.4a (black circles). A genetic algorithm is an iterative approach to solving least-squares problems in which a population of solutions "evolves" toward solutions with the greatest degree of fitness over successive generations. Because of the stochastic nature of this process, in which the initial population is chosen at random, the population tends to arrive at solutions of similar levels of fitness that randomly sample the space randomly. Figure 4.4a shows several example solutions using this approach. We see, as expected, that there is some degree of variability in these solutions. Next, we sought to characterize this variability.

4.4.5 Variability in connection strengths

The most straightforward method for assessing the similarities in the solutions resulting from the GA was to look at the spread in the individual connection strengths. This was done, for each matrix element, by taking the difference between values in the 97.5 and 2.5 percentiles for each connection. The matrix illustrating this spread is shown in Figure 4.4-b. In both the saccadic and optokinetic cases a striking pattern is evident; namely, Connections from short to long time constant and long to short time constant cells cannot change significantly, while connections between cells with similar levels of persistence show a high degree of variability. Furthermore, we see that there is a significant reduction in the variability when the optokinetic responses are included. In fact, the average variability in connection strengths when including both response types is nearly 3 times smaller than when including saccadic responses alone.

While the degree of variability of individual parameters offers some insight into the population of solutions, as illustrated in Figure 4.4-b, assessing the sensitive and insensitive dimensions in solution space provides greater insight into the precise degree of degeneracy in the population. This was done by performing an SVD to identify directions along which the variability was maximal (Methods). The resulting distributions of singular values, which measure the variability along a given direction relative to $W_{\max}$, for the fits involving post-saccadic (top) and post-optokinetic

(bottom) responses are shown in the matrices in Figure 4.4-c. Figure 4.4-d shows an average of all columns within these matrices. As discussed above, when parameter values were unconstrained there exist *N*-2 and *N*-4 insensitive dimension in the fits including saccadic and saccadic+optokinetic responses. Here we find that in the post-saccadic case there are still only 4 dimensions along which the solutions vary by at less than 10% of the maximal weight while there are 27 such dimensions when both responses are included.

4.4.6 Characterization of the general features of the population of solutions

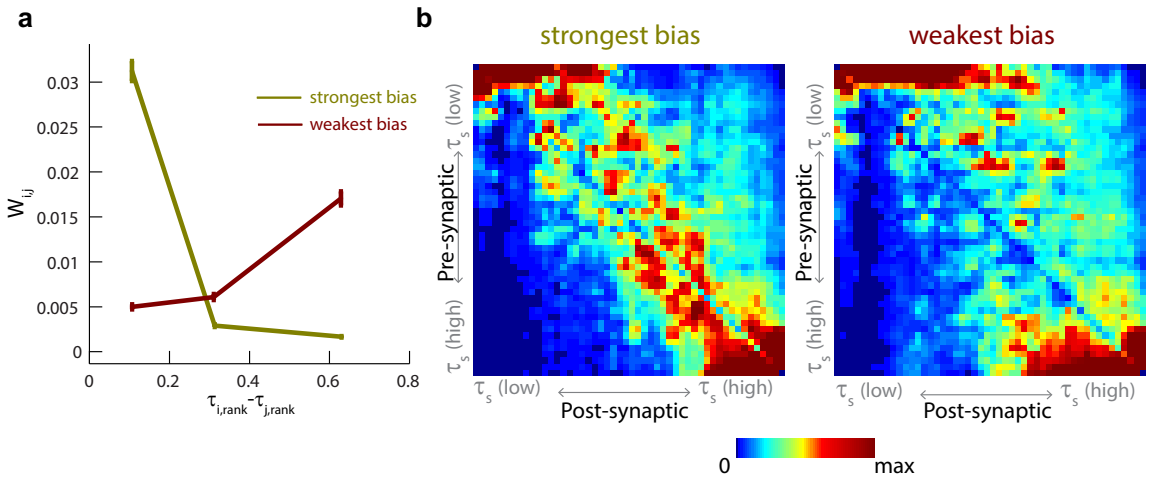


Figure 4.5: Quantifying relationship between connection strength and differences in τ_{sacc}. a.) Data from the two connectivity matrices with the largest negative (green) and largest positive (red) slopes in their connection strength vs. the absolute value in pairwise τ_{sacc} difference. b.) Connectivity matrices corresponding to the plots in a.

In Figure 4.1 we saw that the best-fit, unconstrained solution has some striking structure; namely, that cells connect preferentially to other cells with similar decay time constants. Furthermore, we found that cells also connected preferentially to

nearby cells. We next searched the solutions obtained from the GA process to identify whether the features were universal to the population of solutions.

Figure 4.5 illustrates the wide range of possible values that the connection strength vs. the absolute value in pairwise $\tau_{sacc,rank}$ difference can take. The traces in Figure 4.5a and the corresponding connectivity matrix (Figure 4.5-b) illustrate the connectivity matrices with the largest negative (green) and positive (red) slopes in this relationship. These example matrices were found by sorting all GA solutions based on their slopes. We see that while the network in the right panel of Figure 4.5b tends to have stronger connections between cells with large differences in τ_{sacc}, these connections all tend to take the form of feedforward connections from cells with short time constants to cells with long time constants and not vice versa.

Next, we sorted solutions based on the slope of their connection strength vs. pairwise difference in $\tau_{sacc,rank}$, without taking the absolute value as above in order to search for networks with the feedforward structure. Networks exhibiting a negative slope in this relationship exhibit feedforward connectivity as the connections from cells with short time constants connect more strongly to cells with long time constants. On the other hand networks for which this relationship take on a positive slope would be referred to as feedback because the connections from long to short τ_{sacc} cells are stronger. The results of this search are presented in Figure 4.6, in which

we see that both types of network are present among the GA solutions. Interestingly, in the feedback case, while longer to shorter time constant connections dominate the connectivity, there are no connections between the longest and the shortest time constants cells. This observation is consistent with the top panel of Figure 4.4b which illustrates that such connections are absent from the population of GA solutions.

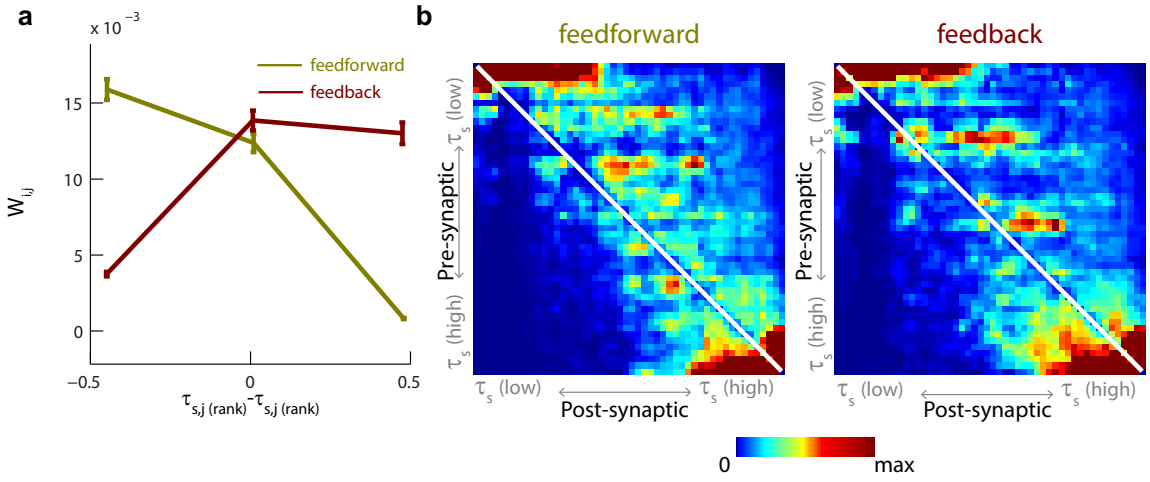


Figure 4.6: Search through GA solutions for post-saccadic responses for feedforward and feedback architectures. Examples of networks that exhibit negative (feedfoward; a, green; b, left) and positive (feedback; a, red; b, right) slopes in their connection strength vs. difference in $\tau_{\text{sacc,rank}}$ plots.

Next characterized the structure of the networks obtained from fitting both post-saccadic and post-optokinetic responses together. In Figure 4.7 we see that the relationship between connection strength and absolute value in the difference in average time constant τ_m is more constrained than the similar relationship shown in Figure 4.5. In fact, in this figure we find that all solutions exhibit a negative slope in

this relationship. This finding illustrates that the integrator network must possess an inverse relationship between connection strength and difference in time constant.

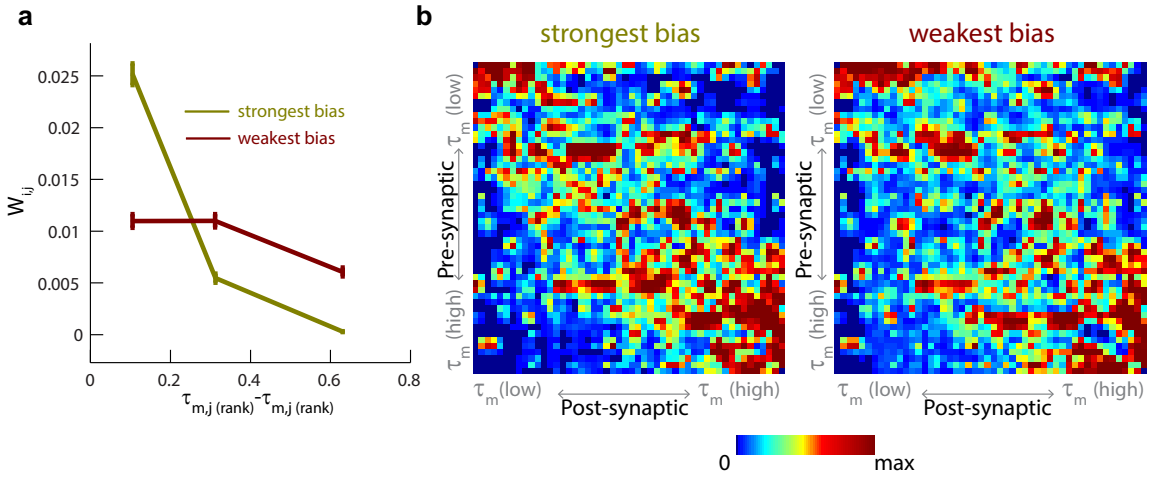


Figure 4.7: Quantifying relationship between connection strength and differences in the average saccadic and optokinetic time constant, τ_m. a.) Data from the two connectivity matrices with the largest negative (green) and largest positive (least negative) (red) slopes in their connection strength vs. the absolute value in pairwise τ_m difference relationships. b.) Connectivity matrices corresponding to the plots in a.

Finally, we search through the GA solutions to both response types for solutions with feedback and feedfoward connectivity. The results of this search are shown in Figure 4.8. Here we see that like the saccadic-only case there are no strong connections from the longest to the shortest time constant cells; and unlike the saccadic-only case, there are no networks with strong feedforward architecture.

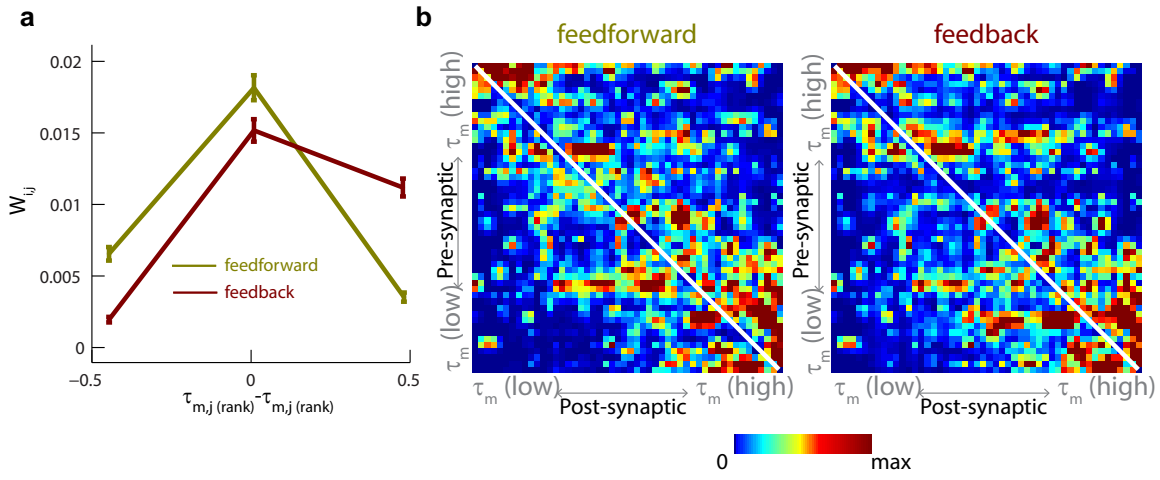


Figure 4.8: Search through GA solutions for both response types for feedforward and feedback architectures. Examples of networks that exhibit negative (feedfoward; a, green; b, left) and positive (feedback; a, red; b, right) slopes in their connection strength vs. difference in τ_m plots.

4.4.7 Searching solutions for spatial structure

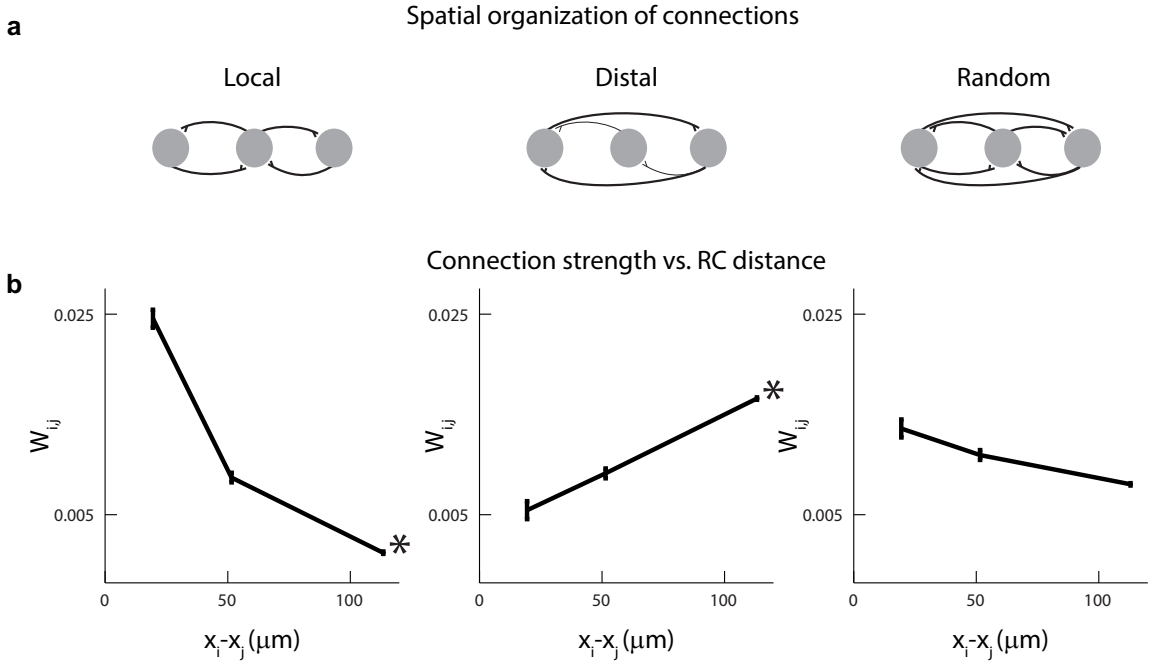


Figure 4.9: Searching GA solutions for spatial organization. a.) Cartoon of spatial organizations discovered in GA solutions. b.) Connection strength vs. distance for local (left) distal (middle) and random (right) connectivity profiles.

Next, we explored the solutions in a similar manner to that described in the previous section, but instead sorted solutions based on their connection vs. spatial separation relationship. This analysis was done on networks fit to both response types. The results of this search are shown Figure 4.9 which illustrates the presence of a much greater diversity in the structures than those shown in Figure 4.7 and Figure 4.8. In fact, we found that networks were present in which connection strength decreased by as much as 100% relative to the average connection strength per every 100 microns of space separating the two cells (Figure 4.9, left), while networks were also present in which this relationship was nearly reversed (Figure 4.9, right) in which connections were stronger between cells that were farther apart.

4.5 Discussion

In this chapter we quantified the degree of variability in the connection strengths in networks fit using the low-dimensional dynamics described in Chapter 3. The variability was quantified using a genetic algorithm to randomly sample from the space of best-fit solutions. Using this approach allowed us to search through the space of solutions for networks with common features and therefore to test whether or not such features were a general quality shared by all networks. We found that using a linear network to model these dynamics necessitates a connectivity structure in which cells with similar average time constants connect preferentially despite the high degree

of variability in the exact connection strengths. On the other hand, we found that the spatial organization of the connections could not be specified as precisely.

Our finding that dynamics recorded following saccades, but not during both behaviors, were consistent with a feedforward pattern of connections illustrates the potential insights to be gained by studying networks under multiple behavioral contexts. Such a result makes intuitive sense when considered in relation to the reversal in persistence gradients shown in Figure 3.5. A network with feedforward connections is capable of generating persistence gradients in one direction, but not in the reverse direction, while a network with a symmetric, chain-like, connectivity is capable in generating gradients whose direction depends on gradients in the inputs.

While the results presented in this chapter depend on the specific model chosen (equation 4.1), the method used to randomly sample the best-fit distribution is generally applicable. For example, a more realistic model of neuronal dynamics involving synaptic non-linearities can be fit using non-linear regression and the space of best-fit solutions could then be randomly sampled using the GA approach outlined here. The GA approach used here represents only one of a number of methods that can be used to randomly sample the space of best-fit solutions. Other approaches in which an exploration of parameter space is guided by the gradient, hessian and knowledge of the hard boundaries could also be used and should provide a similarly large spread in connection weights.

The ambiguity in spatial organization in connection strengths obtained in this chapter motivate the use of perturbative methods as described in Chapter 2 in order to increase the dimensionality of the data. In Chapter 5 we will adapt this approach to help distinguish between the three models shown in Figure 4.9.

Chapter 5 Perturbations of network activity to constrain model fits

5.1 Abstract

In this chapter, we use electrical microstimulations during two-photon calcium imaging to study integrator activity along modes which are not active during behavior. These data are then incorporated into the modeling framework described in Chapter 4 to produce refined estimates of connectivity. We find that each time the stimulating electrode is moved to a new location, a new set of modes is activated in the network, resulting in a reduction in the variability of best-fit connection strengths. This refinement in our estimates of connectivity is associated with the loss of more networks with the "distal" connectivity profiles described in Figure 4.9 from the best-fit distribution than networks with "local" profiles. Simulated experiments suggest that this evolution of the best-fit distribution is consistent with a picture in which the actual connectivity of the integrator network has a "local" structure.

5.2 Introduction

The results of Chapter 4 exposed the difficulty encountered when using regression-based techniques to identify connectivity structures in neural systems with low-dimensional activity. This difficulty arises as specification of all connections within a network requires knowledge of the dynamics supported along all of the network's modes (Figure 1.4). This difficulty is usually avoided in modeling of such dynamics by adding additional "penalty" terms to the cost function, which enables the convergence of the regression algorithm onto one of the many solutions consistent with the data[8,9]. These approaches provide solutions that accurately reproduce the activity along the subset of modes active in the data, but have to essentially guess at what the activity would look like along the remaining dimensions.

In Chapter 2 we introduced a technique in which random perturbations of network activity could be used to provide information about the activity supported by the network along modes not active under natural conditions. When a sufficient number of patterns were excited, the underlying connectivity could be accurately specified (Figures 1.5 and 1.6). One way to achieve activation of distinct patterns of activity is through electrical microstimulation of the whole network. Microstimulation is accomplished by inserting one electrode into the brain area of interest, and another in the fluid surrounding the animal. Neural activity can be evoked when a voltage difference is applied between these two electrodes. This technique has been used to

identify brain regions causally linked to behavior and cognition for over a century[92–94]. Recent work has combined microstimulations with two-photon calcium imaging of cortical network activity[30] to show that the patterns of activity evoked by stimulations were highly dependent on the positioning of the electrode within the network. Here we take advantage of this finding to excite a number of distinct patterns of activity in the integrator by moving the stimulating electrode to multiple different locations.

To provide an estimate of the ability of this approach to distinguish between the "local", "random" and "distal" connectivity structures outlined in Figure 4.9, we first perform simulations in order to understand the way in which the activity of each network differs along modes that were not active during eye movements. With these simulations we find that only a few distinct stimulation patterns are required to begin to distinguish between each of these three models. Next, we confirm experimentally that the pattern of activity evoked by stimulations depends on the positioning of the electrode. Finally, we show that data from stimulations are most consistent with the class of networks in Figure 4.9a in which connections are strongest between neighboring neurons.

5.3 Methods

In Chapter 2 we saw that a randomly-connected 100 neuron connectivity matrix could be accurately reconstructed if the number of distinct patterns of activations was near 7. The aim of this chapter is not to determine the exact connectivity structure of the integrator network, but instead to distinguish between the three classes of structures outlined in Figure 4.9. To accomplish this we first use simulated data to show that given only a small number of stimulations (~3) these network structures can be distinguished by characterizing the way in which the space of best-fit solutions shrinks as stimulations are added to the simulated data sets. Finally, we perform a small number of unique stimulations experimentally to determine which of the structures shown in Figure 4.9 are most consistent with the resulting changes in solution space.

5.3.1 Simulated perturbations

In this chapter all simulations were performed using three of the connectivity matrices obtained using the genetic algorithm described in Chapter 4. These three matrices were chosen on the basis of their weight vs. distance relationship such that the local, distal and random classes were each represented. For each network post-saccadic and/or post-optokinetic responses were simulated by initializing the activity of the network using the initial firing rates determined experimentally, and were

equivalent for each matrix. Additional stimulation patterns were specified by the relationship

$$I_i(t) = \xi e^{-(\mathbf{x_i} - \mathbf{x}_{\text{center}})^2/\sigma^2} \qquad 5.1$$

which gives the current provided to cell *i* at position $\mathbf{x_i}$ with a random amplitude ξ. As the experimental methods used in this chapter to perturb network activity tend to be spatially localized, the currents used in these simulations were also spatially localized about a randomly chosen point within the network $\mathbf{x}_{\text{center}}$. The spatial extent of the stimulation is determined by σ which was chosen to be 50 μm, to be roughly consistent with experimental observations.

Each network was first simulated by providing saccadic inputs only, the resulting activity was then used to calculate a population of best-fit solutions using the genetic algorithm approach described in Chapter 4. Next, network activity was simulated by providing inputs given by equation 5.1. This process was repeated by adding successive patterns of stimulation in a manner similar to that described in Chapter 2. Each time a new pattern of activation was added, the population of best-fit solutions was recalculated. The slope of the connection strength vs. distance relationship was then calculated for the entire distribution of best-fit solutions. This slope was re-calculated each time a new pattern of activation was added. We

compared for the local, distal and random networks how these slopes evolve as successive stimulations are added.

5.3.2 Experimental perturbations

Perturbations of network activity were achieved experimentally by injecting a 500 millisecond triangle pulse of negative current into regions dense in integrator axons and dendrites. Current amplitude was chosen so that pulses evoked small eye movements. The minimal amplitude required to elicit eye movements varied depending on the position on the electrode, but was generally between 1 and 10 μA. Different patterns of stimulation were achieved by either moving the electrode or by changing the current amplitude. For a given experimental session, network activity was first recorded during saccadic eye movements in all planes containing labeled integrator neurons as described in Chapter 3. To avoid distortions of the tissue during the recording session, the electrode was placed into the brain prior to all recording. For a given stimulation pattern (i.e. same electrode position and current amplitude), trains of pluses occurring roughly every 10 seconds were given for between 1 and 2 minutes for each optical plane. This process was then repeated N_{stim}=3 times for each fish.

Integrator neurons were identified based on their correlation with eye position during saccadic eye movements as described in Chapter 3. For each cell, saccadic and N_{stim}=3 different stimulus triggered fluorescence responses were compiled. An SVD

was used to filter the fluorescence responses as described Chapter 3, and the resulting fluorescence responses were then devonvolved yielding firing rate estimates.

Firing rate estimates were then used to estimate a population of best-fit connectivity matrices as in Chapter 4. As with the simulations described above, the space of connectivity matrices was recalculated for each successive stimulation pattern, and the slope of the connection strength vs. distance relationship was determined.

5.4 Results

5.4.1 Network Simulations

Simulations were performed on a local, a distal and a random network (Figure 4.9) obtained from fits to the post-saccadic and post-optokinetic responses of a single fish.

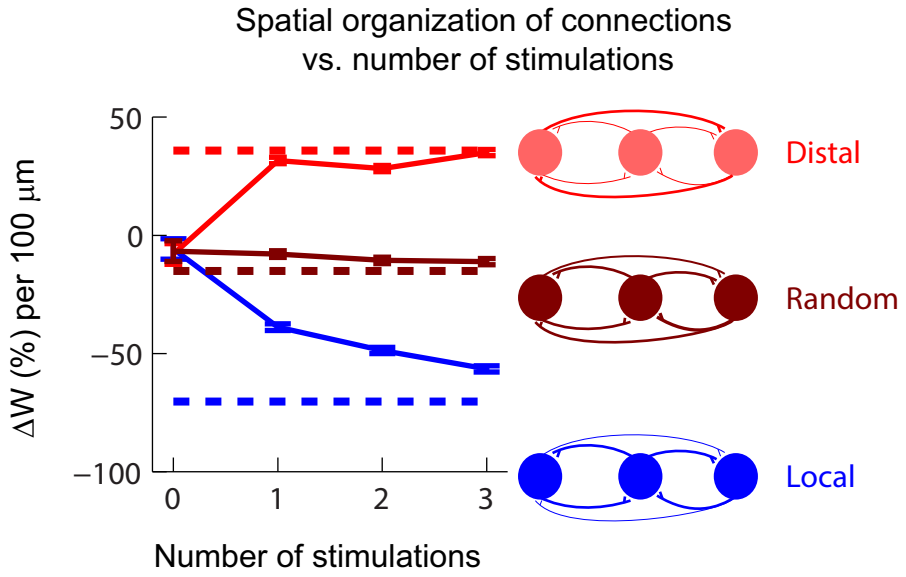


Figure 5.1: Simulations of stimulation experiments. Slopes of the weight vs. distance relationship is plotted as a function of number of independent stimulations for the local (blue), random (maroon) and distal (red) networks identified in Figure 4.9.

We first simulated saccadic responses by initializing each cell in the network with the initial firing rate measured during post-saccadic fixations. As expected, all three networks yielded almost identical patterns of activity during simulations of saccadic responses. After simulating saccadic activity for each network, the distribution of best-fit solutions was then identified. The slope in the relationship between connection weight and spatial separation was computed for weights from all solutions in the distribution. Each time additional patterns of activation are added, a new distribution, and hence a new slope, was calculated. This slope as a function of the number of distinct patterns of non-saccadic stimulation is plotted in . Dotted lines represent the weight vs. distance slope for the actual network. We see from this plot

that as more patterns of activation are added, the solution approaches the correct value of the slope. Based on these results, it should be possible to distinguish between the local, distal and random networks by providing several unique patterns of stimulation to the integrator network.

5.4.2 Experimental Perturbations of Network Activity

Similar to the simulations in the previous section, here we experimentally perturb the network activity in the integrator and characterize the evolution of the weight vs. distance relationship in the distribution of best-fit connectivity matrices to determine which of the local, random or distal connectivity profiles is most consistent with the data.

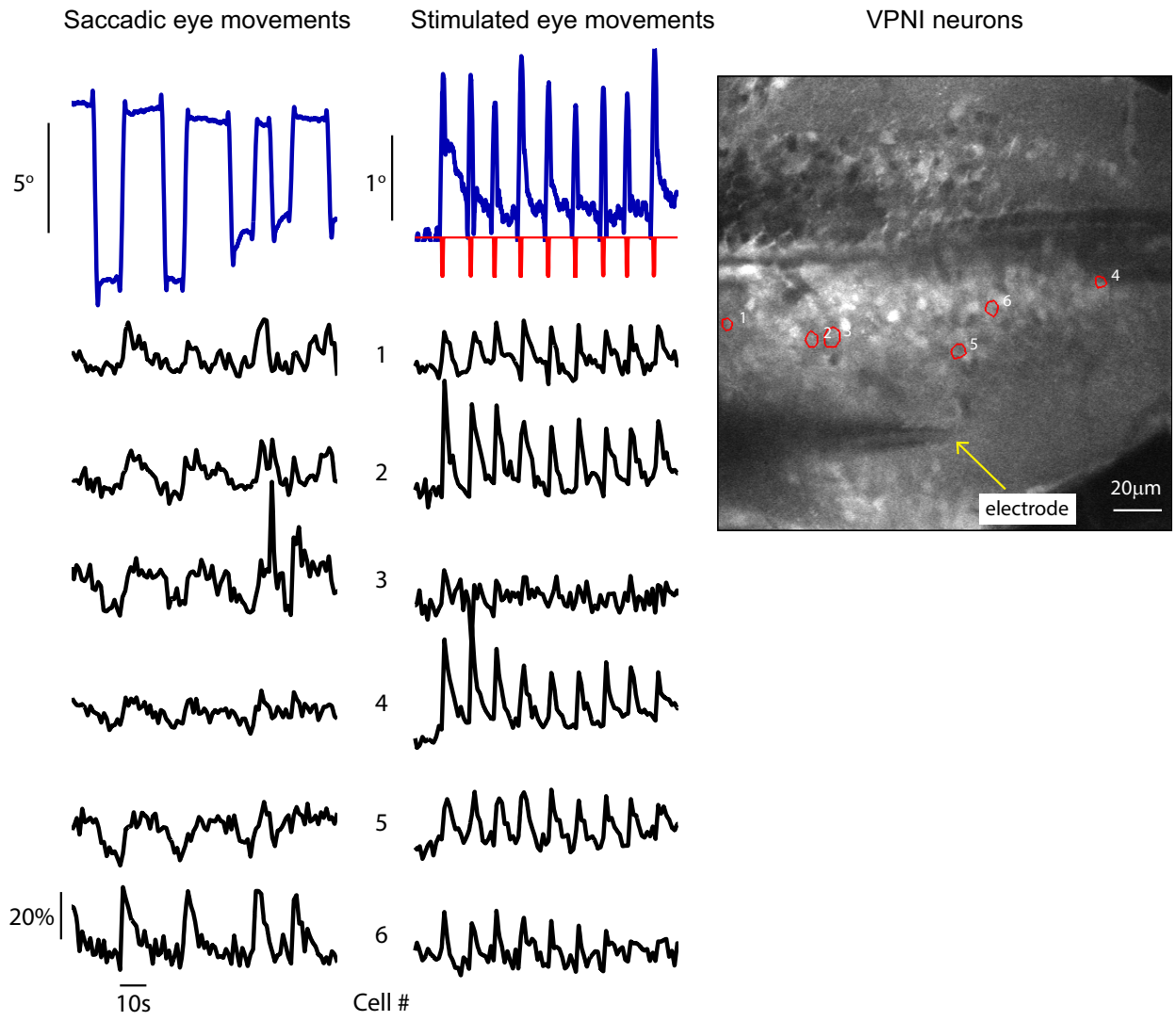


Figure 5.2: Eye position (blue) and activity of integrator neurons (black) are shown for both saccadic and stimulated eye movements. Voltage of stimulating electrode is plotted in red (A.U.).

Brief pulses of current injected into the caudal hindbrain reliably evoked calcium transients in integrator neurons, as well as changes in eye position (Figure 5.2). As illustrated in Figure 5.2**,** calcium transients, and eye position, tended to decay more quickly towards baseline following stimulation than after saccades. In addition, the amplitudes of fluorescence values following stimulation do not appear to be

correlated to the amplitudes following saccades. The differing amplitudes and dynamics during stimulated versus saccadic fluorescence responses suggest that a different set of modes is active following stimulations than following saccades.

To further explore the possibility that different modes are activated by the stimulations, we first calculated stimulation-triggered average fluorescence responses and compared them to the saccade-triggered averages, (Figure 5.3). In Figure 5.3 the responses of 5 cells are plotted during post-saccadic fixations, as well as post-stimulation fixations for 3 distinct stimulation patterns, each created by either moving the stimulating electrode to a new location, or by changing the stimulation amplitude. In this figure, we see that stimulations produce dynamics which differ drastically from post-saccadic responses and that different patterns of stimulation produce distinct patterns of activity. To summarize this finding we performed an SVD on responses following saccades, stimulation pattern 1 and stimulation pattern 2, we then constructed a three-dimensional space from the primary modes from each of these SVDs and projected the 4 responses into this newly constructed space as shown in Figure 5.3. This figure provides a visual illustration of the differences in state-space trajectories produced by different stimulation patterns.

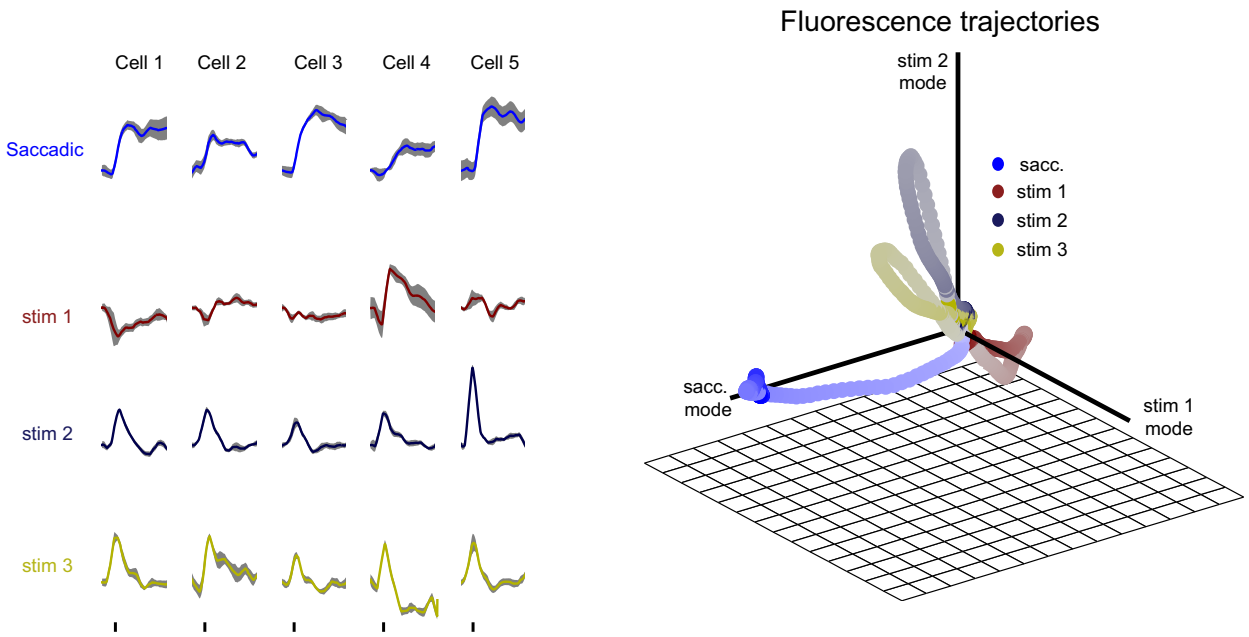


Figure 5.3: Trial-averaged responses (left) are shown for five cells following saccadic (blue) and three distinct stimulation patterns (red, dark blue and yellow). Corresponding state-space trajectories are shown in the right panel.

To quantify the degree to which the dimensions activated by the electrical stimulations we performed a joint SVD on the fluorescence responses for all 3 stimulation patterns and the post-saccadic responses (Figure 5.4 (left)). We found that the responses on both sides of the brain taken together were explained by four modes relative to the noise. In Figure 5.4 (middle), we see that when all responses are combined 16 modes are significant relative to the noise. Figure 5.4 (right) shows that each time a new stimulation pattern is added, more modes are activated.

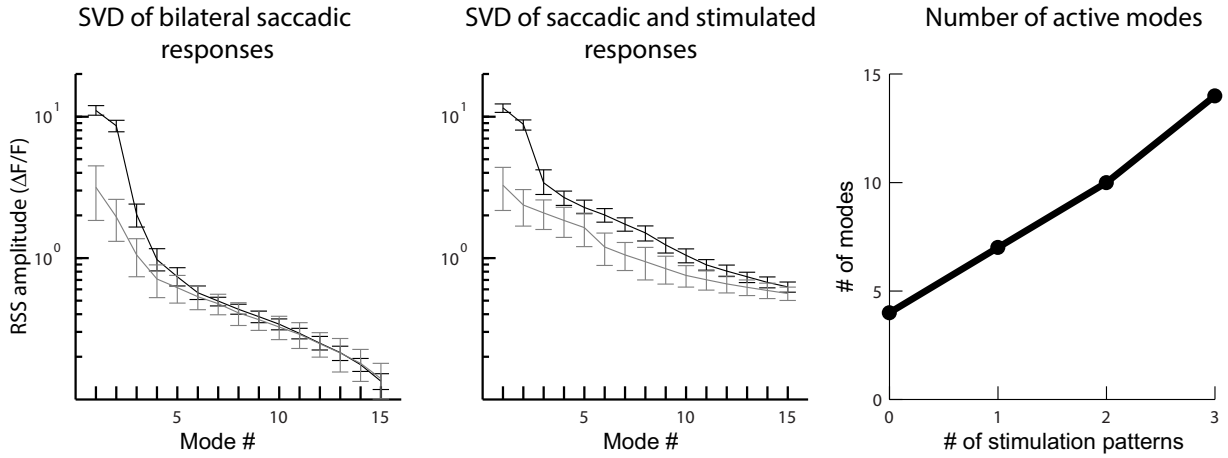


Figure 5.4: SVD of signal (black) and noise (gray) for saccadic (left) and saccadic+stimulated responses (middle). Number of modes whose signal singular values do not overlap with noise singular values are plotted against number of unique stimulation patterns (right).

5.4.3 Post-stimulation dynamics

The increase in dimensionality provided by electrical stimulations will enable us to better constrain our model fits as illustrated in Chapter 2. To incorporate stimulation responses into the modeling framework described in Chapter 4, we first calculated estimates of post-stimulation firing rates using the SVD smoothing and deconvolution procedure outlined in Chapter 3. To reduce noise, all stimulation responses from 3 fish were included combined. In total 347 responses, from 137 cells, were included. As shown Figure 5.5 (right) the SVD yielded three modes that stood out relative to the noise, the components along these modes are shown in Figure 5.5 (middle).

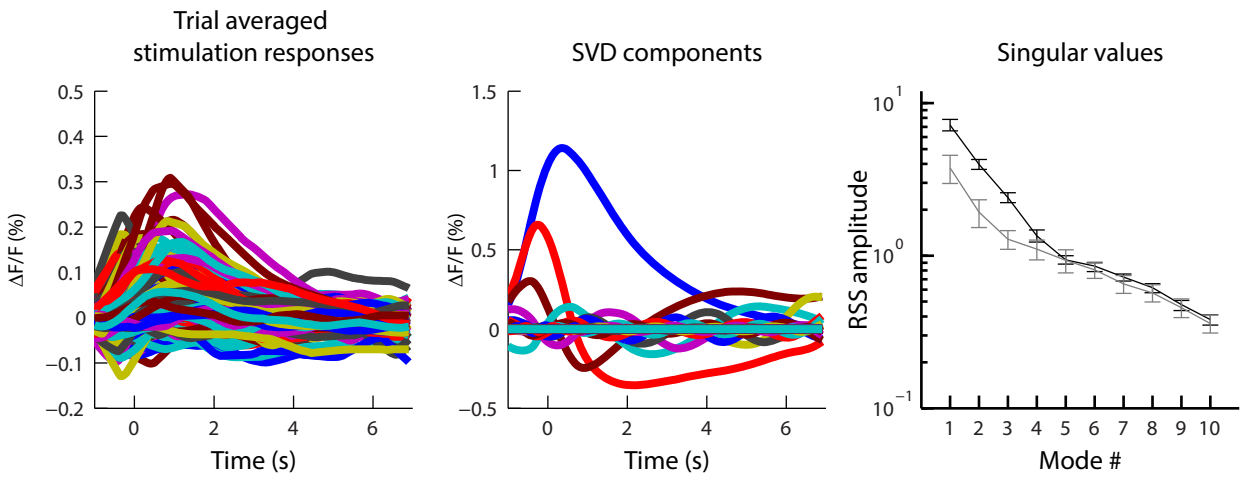
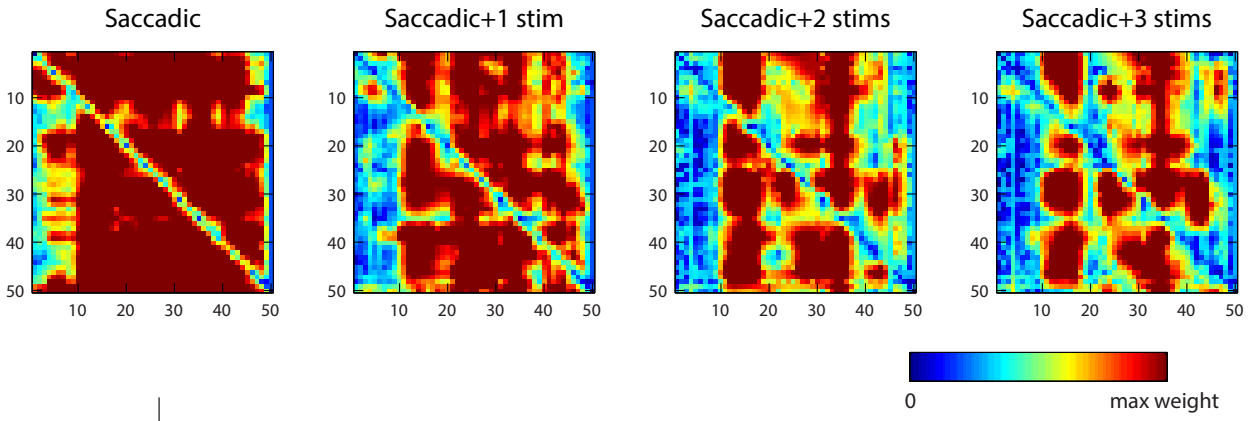


Figure 5.5: All stimulated responses (left), projected components (middle), singular values of signal (black) and noise (gray).

5.4.4 Incorporation of stimulation dynamics into models

Distributions of best-fit solutions were calculated including saccadic responses only, saccadic + stimulation pattern 1, pattern 2…etc. The variability of the connection weights in these distributions are shown in Figure 5.6. In this figure we see that the average variability decreases by more than 25% when three patterns of stimulation are included in the fits.



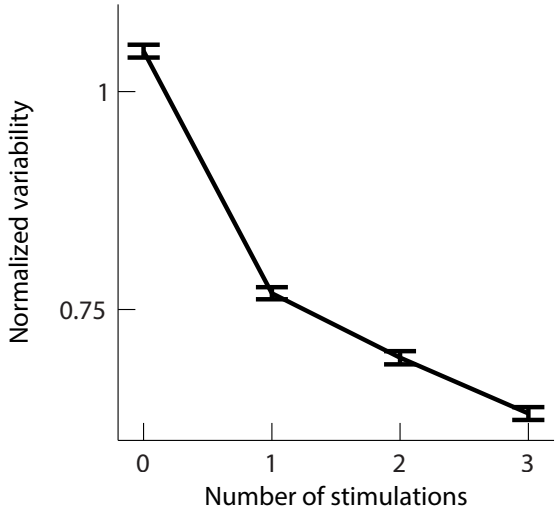


Figure 5.6: Variability in individual connection weights (top). Average variability in connection weigth vs. number of stimulations (bottom).

As in the simulations described in , we calculated the slopes of the connection weight vs. distance relationship as a function of the number of stimulation patterns. The results of this analysis are shown in Figure 5.7 in which we see that each time a new pattern of stimulation is added, the slope becomes more negative. As shown in , this type of behavior is consistent with a network in which the connections are strongest between neighboring neurons. Intuitively the increasingly negative slopes can be thought of as arising from the loss of solutions with strong connections between far-away cell pairs.

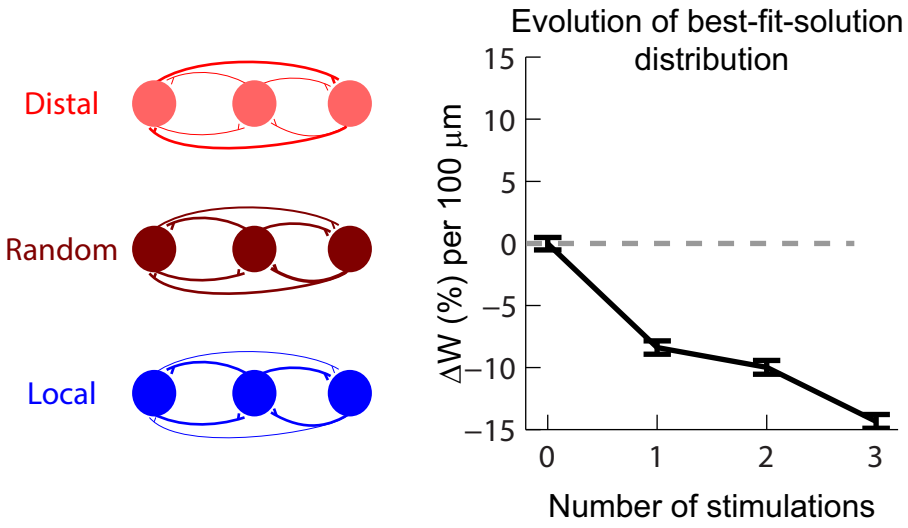


Figure 5.7: Slope of weight vs. distance relationship as a function of number of stimulations.

5.5 Discussion

In this chapter, we used electrical stimulations to perturb network activity into dimensions in state space that were unexplored during behavior (Figure 5.2). These data were then used to refine estimates of connectivity. This refinement was quantified using the framework described in Chapter 4. Each time a new pattern of stimulation was added to the fits, the variability in connection weights decreased (Figure 5.7). In other words, the number of distinct matrices consistent with the data decreased. We found that a greater proportion of networks with strong connections between distal pairs of neurons were lost relative to networks with strong connections between nearby neurons. This evolution in the distribution of best-fit solutions, as supported by the simulation shown in , are consistent with a picture in which the true

underlying connectivity matrix is organized such that connections are strongest between neighboring neurons.

As discussed in Chapter 4, the results presented in this chapter depend on the particular form of the equation used to model the network activity. However, the combined experimental and theoretical tools used to refine estimates of connectivity will be capable of providing refined estimates of connectivity independent of the model used.

In an attempt to identify anatomical correlates of the functional observation made in this chapter, namely stronger connections between nearby neurons, we have recently begun an effort to use electron microscopy to image connections between all functionally-identified integrator neurons in a single brain. This effort should be able elucidate the long-standing question how these neurons physically connect (i.e. gap junctions, dendro-dendritic or axo-dendritic synapses). Furthermore, these data should help further refine the bounds on the slope of the connection strength vs. inter-cellular distance relationship.

Chapter 6 Conclusions

6.1 Overview

In this thesis we have developed an experimental framework for investigating the functional interactions between neurons in a microcircuit. The basis for this framework is the idea that the state of a neural circuit can be fully described as a trajectory through an *N*-dimensional state space; and the allowable trajectories are determined by the underlying connectivity of the circuit. Because of this idea, the accuracy of estimates of connectivity increase as the activity of the system is observed along an increasing number of dimensions. With these implications in mind, we designed a set of experiments with the objective of maximizing the dimensionality of our data, enabling us to specify the most likely functional and spatial organizational motifs underlying the generation of persistent firing in the integrator. These results provide an instructive example of the additional insights to be gained from the consideration of quantitative network implications in the design of experiments in neuroscience.

As described in Chapter 1, the connectivity in oculomotor integrator has traditionally been modeled as a line attractor[15], motivated by the approximately linear

relationship between eye position and firing rate[24]. In these models, persistent activity is present along a single mode in the network, while the *N*-1 remaining modes support leaky dynamics. In Chapter 3 we tested for the presence of multiple modes supporting non-trivial dynamics. We hypothesized that different types of eye movements, each operating at different speeds—and therefore utilizing different proportions of fast and slow-twitch fibers—engage different modes in the network. Confirming this idea, we found that different modes supporting activity persisting for multiple seconds were engaged by saccadic and optokinetic eye movements. Not only do these results provide insights into how networks are able to effectively control different behaviors, but, as shown in Chapter 4, also provide refined estimates of the underlying connectivity in the network. In Chapter 5 we show that combining a small number of random perturbations of network activity with recordings of population dynamics are capable of producing accurate estimates of the spatial-connectivity motif present in the integrator network.

While this method could potentially be used to determine the connectivity of microcircuits in other brain regions, there are potential difficulties not encountered in the integrator that must be taken into consideration. Estimates of connectivity in the integrator are simpler than those of other networks because the dynamics of the network during fixations are produced almost entirely by intra-network interactions. Estimates of connectivity in other brain regions, such as cortex, will prove to be more

complicated. While distinct cortical regions tend to be responsible for different computations, these regions tend to be highly interconnected[95,96]. This interconnectedness makes estimates of connectivity using the methods described in this thesis more difficult. However, the same general ideas put forth in the preceding five chapters could, in principle, lead to the same level of insight regarding connectivity in cortical microcircuits. The main technical challenge to overcome in applying these state space techniques to studying the connectivity in regions that are highly interconnected with other regions is to be able to record from each of these regions during perturbations either simultaneously or sequentially as done in Chapter 5. Given the intense effort currently being expended to dramatically increase the number of cells that can be recorded from in a given experimental session[3–5], such an approach should be achievable in the near future.

References

1. Stosiek, C., Garaschuk, O., Holthoff, K. & Konnerth, A. In vivo two-photon calcium imaging of neuronal networks. *Proc. Natl. Acad. Sci.* **100,** 7319–7324 (2003).
2. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* **261,** 1055–1058 (1993).
3. Alivisatos, A. P. *et al.* The Brain Activity Map Project and the Challenge of Functional Connectomics. *Neuron* **74,** 970–974 (2012).
4. Alivisatos, A. P. *et al.* The Brain Activity Map. *Science* **339,** 1284–1285 (2013).
5. Alivisatos, A. P. *et al.* Nanotools for neuroscience and brain activity mapping. *ACS Nano* **7,** 1850–1866 (2013).
6. Bock, D. D. *et al.* Network anatomy and in vivo physiology of visual cortical neurons. *Nature* **471,** 177–182 (2011).
7. Briggman, K. L., Helmstaedter, M. & Denk, W. Wiring specificity in the direction-selectivity circuit of the retina. *Nature* **471,** 183–188 (2011).
8. Fisher, D., Olasagasti, I., Aksay, E., Tank, D. W. & Goldman, M. A Modeling Framework for Deriving the Structural and Functional Architecture of a Short-Term Memory Microcircuit. *Neuron* **79,** 987–1000 (2013).

9. Machens, C. K., Romo, R. & Brody, C. D. Functional, But Not Anatomical, Separation of 'What' and 'When' in Prefrontal Cortex. *J Neurosci* **30,** 350–360 (2010).
10. Bernacchia, A., Seo, H., Lee, D. & Wang, X.-J. A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* **14,** 366–372 (2011).
11. Miri, A. *et al.* Spatial gradients and multidimensional dynamics in a neural integrator circuit. *Nat. Neurosci.* **14,** 1150–1159 (2011).
12. Gutenkunst, R. N. *et al.* Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* **3,** 1871–1878 (2007).
13. Hopfield, J. J. Neurons with Graded Response Have Collective Computational Properties like Those of Two-State Neurons. *Proc. Natl. Acad. Sci. U. S. A.* **81,** 3088–3092 (1984).
14. Hopfield, J. & Tank, D. Computing with neural circuits: a model. *Science* **233,** 625–633 (1986).
15. Seung, H. S. How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13339–13344 (1996).
16. Briggman, K. L., Abarbanel, H. D. I. & Kristan, W. B. Optical Imaging of Neuronal Populations During Decision-Making. *Science* **307,** 896–901 (2005).
17. Niessing, J. & Friedrich, R. W. Olfactory pattern classification by discrete neuronal network states. *Nature* **465,** 47–52 (2010).

18. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* (2012). doi:10.1038/nature11129

19. Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484,** 62–68 (2012).

20. Brody, C. D., Romo, R. & Kepecs, A. Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Curr. Opin. Neurobiol.* **13,** 204–211 (2003).

21. Major, G. & Tank, D. Persistent neural activity: prevalence and mechanisms. *Curr. Opin. Neurobiol.* **14,** 675–684 (2004).

22. Ganguli, S. *et al.* One-Dimensional Dynamics of Attention and Decision Making in LIP. *Neuron* **58,** 15–25 (2008).

23. Burak, Y. & Fiete, I. R. Accurate Path Integration in Continuous Attractor Network Models of Grid Cells. *PLoS Comput Biol* **5,** e1000291 (2009).

24. Aksay, E., Baker, R., Seung, H. S. & Tank, D. W. Anatomy and Discharge Properties of Pre-Motor Neurons in the Goldfish Medulla That Have Eye-Position Signals During Fixations. *J Neurophysiol* **84,** 1035–1049 (2000).

25. Aksay, E., Gamkrelidze, G., Seung, H. S., Baker, R. & Tank, D. W. In vivo intracellular recording and perturbation of persistent activity in a neural integrator. *Nat Neurosci* **4,** 184–193 (2001).

26. Aksay, E., Baker, R., Seung, H. S. & Tank, D. W. Correlated discharge among cell pairs within the oculomotor horizontal velocity-to-position integrator. *J. Neurosci.* **23,** 10852–10858 (2003).

27. Aksay, E. *et al.* Functional dissection of circuitry in a neural integrator. *Nat Neurosci* **10,** 494–504 (2007).

28. Tsodyks, M. V. & Markram, H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci.* **94,** 719–723 (1997).

29. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat Neurosci* **8,** 1263–1268 (2005).

30. Histed, M. H., Bonin, V. & Reid, R. C. Direct Activation of Sparse, Distributed Populations of Cortical Neurons by Electrical Microstimulation. *Neuron* **63,** 508–522 (2009).

31. Baddeley, A. The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* **4,** 417–423 (2000).

32. Fuster, J. M. The Prefrontal Cortex—An Update: Time Is of the Essence. *Neuron* **30,** 319–333 (2001).

33. Major, G. & Tank, D. Persistent neural activity: prevalence and mechanisms. *Curr. Opin. Neurobiol.* **14,** 675–684 (2004).

34. Fuster, J. M. & Alexander, G. E. Neuron Activity Related to Short-Term Memory. *Science* **173,** 652–654 (1971).

35. Romo, R., Brody, C. D., Hernandez, A. & Lemus, L. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* **399,** 470–473 (1999).

36. Shadlen, M. N. & Newsome, W. T. Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. *J. Neurophysiol.* **86,** 1916–1936 (2001).

37. Robinson, D. A. Eye Movement Control in Primates. *Science* **161,** 1219–1224 (1968).

38. Cohen, B. & Komatsuzaki, A. Eye movements induced by stimulation of the pontine reticular formation: Evidence for integration in oculomotor pathways. *Exp. Neurol.* **36,** 101–117 (1972).

39. Lopez-Barneo, J., Darlot, C., Berthoz, A. & Baker, R. Neuronal activity in prepositus nucleus correlated with eye movement in the alert cat. *J. Neurophysiol.* **47,** 329–352 (1982).

40. McFarland, J. L. & Fuchs, A. F. Discharge Patterns in Nucleus Prepositus Hypoglossi and Adjacent Medial Vestibular Nucleus During Horizontal Eye Movement in Behaving Macaques. *J. Neurophysiol.* **68,** 319–332 (1992).

41. Kaneko, C. R. S. Eye Movement Deficits Following Ibotenic Acid Lesions of the Nucleus Prepositus Hypoglossi in Monkeys II. Pursuit, Vestibular, and Optokinetic Responses. *J. Neurophysiol.* **81,** 668–681 (1999).

42. McCrea, R. A. & Baker, R. Anatomical connections of the nucleus prepositus of the cat. *J. Comp. Neurol.* **237,** 377–407 (1985).

43. Prevosto, V., Graf, W. & Ugolini, G. Posterior parietal cortex areas MIP and LIPv receive eye position and velocity inputs via ascending preposito-thalamo-cortical pathways. *Eur. J. Neurosci.* **30,** 1151–1161 (2009).

44. Seung, H. S., Lee, D. D., Reis, B. Y. & Tank, D. W. Stability of the Memory of Eye Position in a Recurrent Network of Conductance-Based Model Neurons. *Neuron* **26,** 259–271 (2000).

45. Goldman, M. S., Levine, J. H., Major, G., Tank, D. W. & Seung, H. S. Robust Persistent Neural Activity in a Model Integrator with Multiple Hysteretic Dendrites per Neuron. *Cereb Cortex* **13,** 1185–1195 (2003).

46. Goldman, M., Compte, A. & Wang, X.-J. in *New Encycl. Neurosci.* **6,** 165–178 (Oxford: Academic Press, 2009).

47. Aksay, E. *et al.* History dependence of rate covariation between neurons during persistent activity in an oculomotor integrator. *Cereb. Cortex N. Y. N 1991* **13,** 1173–1184 (2003).

48. Joshua, M., Medina, J. F. & Lisberger, S. G. Diversity of Neural Responses in the Brainstem during Smooth Pursuit Eye Movements Constrains the Circuit Mechanisms of Neural Integration. *J. Neurosci.* **33,** 6633–6647 (2013).

49. Cannon, S. C. & Robinson, D. A. Loss of the neural integrator of the oculomotor system from brain stem lesions in monkey. *J. Neurophysiol.* **57,** 1383–1409 (1987).

50. Miri, A., Daie, K., Burdine, R. D., Aksay, E. & Tank, D. W. Regression-Based Identification of Behavior-Encoding Neurons During Large-Scale Optical Imaging of Neural Activity at Cellular Resolution. *J. Neurophysiol.* **105,** 964–980 (2011).

51. Scudder, C. A., Kaneko, C. S. & Fuchs, A. F. The brainstem burst generator for saccadic eye movements: a modern synthesis. *Exp. Brain Res. Exp. Hirnforsch. Expérimentation Cérébrale* **142,** 439–462 (2002).

52. Mustari, M. J. & Fuchs, A. F. Discharge patterns of neurons in the pretectal nucleus of the optic tract (NOT) in the behaving primate. *J. Neurophysiol.* **64,** 77–90 (1990).

53. Mustari, M. J., Fuchs, A. F., Kaneko, C. R. & Robinson, F. R. Anatomical connections of the primate pretectal nucleus of the optic tract. *J. Comp. Neurol.* **349,** 111–128 (1994).

54. Helmchen, F. & Tank, D. W. in *Imaging Neurosci. Lab. Man.* (Helmchen, F. & Konnerth, A.) 355–368 (University of Zurich, 2011).

55. Cui, H. & Andersen, R. A. Posterior Parietal Cortex Encodes Autonomously Selected Motor Plans. *Neuron* **56,** 552–559 (2007).

56. Stokes, M. G. *et al.* Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron* **78,** 364–375 (2013).

57. Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497,** 585–590 (2013).

58. Rao, S. C., Rainer, G. & Miller, E. K. Integration of What and Where in the Primate Prefrontal Cortex. *Science* **276,** 821–824 (1997).

59. Cullen, K. E. The vestibular system: multimodal integration and encoding of self-motion for motor control. *Trends Neurosci.* **35,** 185–196 (2012).

60. Fetsch, C. R., Pouget, A., DeAngelis, G. C. & Angelaki, D. E. Neural correlates of reliability-based cue weighting during multisensory integration. *Nat. Neurosci.* **15,** 146–154 (2012).

61. Pouget, A., Deneve, S. & Duhamel, J.-R. A computational perspective on the neural basis of multisensory spatial representations. *Nat. Rev. Neurosci.* **3,** 741–747 (2002).

62. Dieringer, N. & Precht, W. Functional organization of eye velocity and eye position signals in abducens motoneurons of the frog. *J. Comp. Physiol. A* **158,** 179–194 (1986).

63. Straka, H. & Dieringer, N. Basic organization principles of the VOR: lessons from frogs. *Prog. Neurobiol.* **73,** 259–309 (2004).

64. Büttner-Ennever, J. A., Horn, A. K., Scherberger, H. & D'Ascanio, P. Motoneurons of twitch and nontwitch extraocular muscle fibers in the abducens, trochlear, and oculomotor nuclei of monkeys. *J. Comp. Neurol.* **438,** 318–335 (2001).

65. Büttner-Ennever, J. A. in *Neuroanat. Oculomot. Syst.* **Volume 151,** 95–125 (Elsevier, 2006).

66. Sklavos, S., Dimitrova, D. M., Goldberg, S. J., Porrill, J. & Dean, P. Long Time-Constant Behavior of the Oculomotor Plant in Barbiturate-Anesthetized Primate. *J. Neurophysiol.* **95,** 774–782 (2006).

67. Sommer, M. A. & Wurtz, R. H. Brain Circuits for the Internal Monitoring of Movements. *Annu. Rev. Neurosci.* **31,** 317–338 (2008).

68. Lisberger, S. G. Internal models of eye movement in the floccular complex of the monkey cerebellum. *Neuroscience* **162,** 763–776 (2009).

69. Major, G. *et al.* Plasticity and tuning by visual feedback of the stability of a neural integrator. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 7739–7744 (2004).

70. Major, G., Baker, R., Aksay, E., Seung, H. S. & Tank, D. W. Plasticity and Tuning of the Time Course of Analog Persistent Firing in a Neural Integrator. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 7745–7750 (2004).

71. MacNeil, D. & Eliasmith, C. Fine-Tuning and the Stability of Recurrent Neural Networks. *PLoS ONE* **6,** e22885 (2011).

72. Arnold, D. B. & Robinson, D. A. A learning network model of the neural integrator of the oculomotor system. *Biol. Cybern.* **64,** 447–454 (1991).

73. Shen, L. Neural integration by short term potentiation. *Biol. Cybern.* **61,** 319–325 (1989).

74. Mongillo, G., Barak, O. & Tsodyks, M. Synaptic Theory of Working Memory. *Science* **319,** 1543–1546 (2008).

75. Koulakov, A. A., Raghavachari, S., Kepecs, A. & Lisman, J. E. Model for a robust neural integrator. *Nat. Neurosci.* **5,** 775–782 (2002).

76. Dickinson, P. S., Mecsas, C. & Marder, E. Neuropeptide fusion of two motor-pattern generator circuits. *Nature* **344,** 155–158 (1990).

77. Goldman, M. S., Golowasch, J., Marder, E. & Abbott, L. F. Global Structure, Robustness, and Modulation of Neuronal Models. *J. Neurosci.* **21,** 5229–5238 (2001).

78. Heys, J. G. & Hasselmo, M. E. Neuromodulation of Ih in Layer II Medial Entorhinal Cortex Stellate Cells: A Voltage-Clamp Study. *J. Neurosci.* **32,** 9066–9072 (2012).

79. Delgado-García, J. M., Yajeya, J. & Navarro-López, J. de D. in *Prog. Brain Res.* (S. Martinez-Conde, S. L. M.) **154, Part A,** 211–224 (Elsevier, 2006).

80. Giocomo, L. M. & Hasselmo, M. E. Knock-Out of HCN1 Subunit Flattens Dorsal–Ventral Frequency Gradient of Medial Entorhinal Neurons in Adult Mice. *J. Neurosci.* **29,** 7625–7630 (2009).

81. Fuchs, A. F. & Kimm, J. Unit activity in vestibular nucleus of the alert monkey during horizontal angular acceleration and eye movement. *J. Neurophysiol.* **38,** 1140–1161 (1975).

82. Zucker, R. S. & Regehr, W. G. Short-Term Synaptic Plasticity. *Annu. Rev. Physiol.* **64,** 355–405 (2002).

83. Gulledge, A. T. & Jaffe, D. B. Multiple Effects of Dopamine on Layer V Pyramidal Cell Excitability in Rat Prefrontal Cortex. *J. Neurophysiol.* **86,** 586–595 (2001).

84. Smetters, D., Majewska, A. & Yuste, R. Detecting Action Potentials in Neuronal Populations with Calcium Imaging. *Methods* **18,** 215–221 (1999).

85. Brustein, E., Marandi, N., Kovalchuk, Y., Drapeau, P. & Konnerth, A. 'In vivo' monitoring of neuronal network activity in zebrafish by two-photon Ca 2+ imaging. *Pflüg. Arch. Eur. J. Physiol.* **446,** 766–773 (2003).

86. Beck, J. C., Gilland, E., Tank, D. W. & Baker, R. Quantifying the Ontogeny of Optokinetic and Vestibuloocular Behaviors in Zebrafish, Medaka, and Goldfish. *J Neurophysiol* **92,** 3546–3561 (2004).

87. Yaksi, E. & Friedrich, R. W. Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca2+ imaging. *Nat Meth* **3,** 377–383 (2006).

88. Machens, C. K., Romo, R. & Brody, C. D. Functional, But Not Anatomical, Separation of 'What' and 'When' in Prefrontal Cortex. *J. Neurosci. Off. J. Soc. Neurosci.* **30,** 350–360 (2010).

89. Demer, J. L. & Robinson, D. A. Different time constants for optokinetic and vestibular nystagmus with a single velocity-storage element. *Brain Res.* **276,** 173–177 (1983).

90. Beck, J. C., Rothnie, P., Straka, H., Wearne, S. L. & Baker, R. Precerebellar Hindbrain Neurons Encoding Eye Velocity During Vestibular and Optokinetic Behavior in the Goldfish. *J Neurophysiol* **96,** 1370–1382 (2006).

91. Ma, L.-H., Punnamoottil, B., Rinkwitz, S. & Baker, R. Mosaic hoxb4a Neuronal Pleiotropism in Zebrafish Caudal Hindbrain. *PLoS ONE* **4,** e5944 (2009).

92. Fritsch, G. & Hitzig, E. Ueber die elektrische Erregbarkeit des Gros- shirns. Archiv fur Anatomie, Physiologie und wissenschaftliche. *Medizin* **37,** 300–339 (1870).

93. Penfield, W. Ferrier Lecture: Some Observations on the Cerebral Cortex of Man. *Proc. R. Soc. Lond. B Biol. Sci.* **134,** 329–347 (1947).

94. Salzman, C. D., Britten, K. H. & Newsome, W. T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature* **346,** 174–177 (1990).

95. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex N. Y. N 1991* **1,** 1–47 (1991).

96. Mesulam, M. M. From sensation to cognition. *Brain* **121,** 1013–1052 (1998).